PB87-142337

Evaluation of Short-Term Long-Range Transport Models. Volume 1 Analysis Procedures and Results

Argonne National Lab., IL

Prepared for

Environmental Protection Agency
Research Triangle Park, NC

Oct 86

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-450/4-86-016a	2.	3. RECIPIENT'S ACCESSION NO. PB87 142337/AS
4. TITLE AND SUBTITLE Evaluation of Short-Term Long-Range Transport Models--Volume I Analysis Procedures and Results		5. REPORT DATE date of preparation October 1986 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) A. J. Policastro, M. Wastag, L. Coke, R. A. Carhart, and W. E. Dunn		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Argonne National Laboratory Argonne, Illinois 60439 U. of Ill. Chicago, Ill 60680 U. of Illinois Urbana, Ill. 61801		10. PROGRAM ELEMENT NO. B24A2F 11. CONTRACT/GRANT NO. DW89930807
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Office of Air Quality Planning and Standards Research Triangle Park, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED Final report 14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

Eight short-term long-range transport models (MESOPUFF, MESOPLUME, MSPUFF, MESOPUFF II, MTDDIS, ARRPA, RADM, and RTM-II) have been evaluated with field data from two data bases involving tracer releases. The primary quantitative means of evaluating model performance was the use of the American Meteorological Society statistics. Supplementary measures included the use of isopleth plots of ground-level concentrations, scatter plots, cumulative frequency distributions and frequency histograms of residuals. General features of the model performance included: (a) spatial offset of predicted and observed patterns, (b) a time difference between the arrival of the predicted and observed plumes at a particular receptor, and (c) an angular offset of as much as 20-45 degrees between predicted and observed plumes. The models also tended to underpredict horizontal spreading at ground level, along with overprediction of plume concentrations. As a result, predicted concentrations correlated poorly with concentrations observed at the same time and place. However, statistical comparisons of the peak values predicted by the models were significantly better. For example, the highest 25 averaged predictions and highest 25 averaged observations (unpaired in location and time) were within a factor or two of each other for six of the eight models tested (MESOPUFF, MESOPLUME, MESOPUFF II, MTDDIS, ARRPA, and RTM-II).

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution Long Range Transport Models Meteorology Model Evaluation Statistics		

18. DISTRIBUTION STATEMENT Release unlimited	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 2214
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

PB87-142337

United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-86-016a
October 1986

Air

EPA

Evaluation of Short-Term Long-Range Transport Models

Volume I. Analysis Procedures and Results

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA. 22161

EPA-450/4-86-016a

Evaluation of Short-Term Long-Range Transport Models
Volume I. Analysis Procedures and Results

by
A. J. Policastro, M. Wastag, L. Coke
Argonne National Laboratory
Argonne, Illinois 60439

R. A. Carhart
University of Illinois
Chicago, Illinois 60680

W. E. Dunn
University of Illinois
Urbana, Illinois 61801

IAG No. DW 89930807

EPA Project Officer: Norman C. Possiel

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Radiation
Office of Air Quality Planning and Standards
Source Receptor Analysis Branch
Research Triangle Park, NC 27711

October 1986

1.a

Disclaimer

This report has been reviewed by The Office of Air Quality Planning and Standards, U. S. Environmental Protection Agency, and has been approved for publication. Mention of trade names or commercial products is not intended to constitute endorsement or recommendation for use.

ABSTRACT

Eight short-term long-range transport models have been evaluated with field data from two data bases involving tracer releases. The models tested were MESOPUFF, MESOPLUME, MSPUFF, MESOPUFF II, MTDDIS, ARRPA, RADM, and RTM-II. The Oklahoma data base encompasses two separate experiments with measurements taken at 100 km and 600 km arcs downwind of a three-hour perfluorocarbon tracer release. Time averages of 45 minutes for the 100 km arc and 3 hours for the 600 km arc were obtained in the data. The spatial detail along these arcs was very good providing an opportunity to evaluate such predicted plume features as transport time, horizontal plume spreading, and predicted/observed pattern offsets. The Savannah River Plant data base encompasses 15 experiments with measurements taken over 2-5 days at distances 28-144 km downwind. Ten-hour time averages of the krypton-85 tracer were taken at ground level. This data base involved a longer period of record and meteorology from four seasons, yet included a fewer number of fixed samplers than the Oklahoma experiments.

Model performance was evaluated by graphical and statistical methods. The primary means of evaluating the performance of the models was the use of the American Meteorological Society (AMS) statistics. These statistics provided quantitative measures of model performance. Supplementary measures included the use of isopleth plots of ground-level concentrations, scatter plots, cumulative frequency distributions and frequency histograms of residuals.

Model performance was generally consistent between the two data bases. General features of the predicted ground patterns included: (a) spatial offset of predicted and observed patterns, (b) a time difference between the arrival of the predicted and observed plumes at a particular receptor, (c) a definite angular offset of predicted and observed plumes generally due to specific assumptions made in the meteorological preprocessor. This angular offset is frequently as much as 20-45 degrees. The ARRPA and MTDDIS models appear to provide the most accurate trajectories of all eight models. The ARRPA model employs the Boundary Layer Model of the National Weather Service for its meteorological preprocessor, and the MTDDIS model employs wind directions at the effective height of release. However, it should be stated that the ARRPA

model could only be tested with the Oklahoma data base and the MTDDIS model was applicable only to the Savannah River Plant data base. Further model testing would be needed in order to generalize the above findings on plume trajectory predictions.

The models also have a definite tendency to underpredict horizontal spreading at ground level, along with a concomitant overprediction of plume concentrations. This behavior is typical of the performance of most models. The underprediction of horizontal spreading is usually due to the use of the Turner curves at distances (50-100 km) far beyond their intended use.

The spatial and temporal offsets of the predicted and observed plumes lead to predicted concentrations that correlate poorly with concentrations observed at the same time and place. Values for standard deviation of residuals, root mean square error, and absolute average residual are larger than the average observed concentration for all eight models. However, statistical comparisons of the peak values predicted by the models were significantly better. For example, the highest 25 averaged predictions and highest 25 averaged observations (unpaired in location and time) were within a factor of two of each other for six of the eight models tested (MESOPUFF, MESOPLUME, MESOPUFF II, MTDDIS, ARRPA, and RTM-II). Again, the ARRPA and MTDDIS models were tested with only one data base. The observed tendency to overpredict peak concentrations errs on the conservative side for regulatory applications. However, this overprediction must be weighed against the general tendency of the models to underpredict horizontal spreading and to predict a pattern that is spatially offset from the data.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Norman Possiel and Mr. Joseph Tikvart of the U.S. EPA for their support and direction of this model evaluation work. Their structuring of the project to permit periodic reviews of project accomplishments by the model developers was important to the success of this work.

The authors are also appreciative of the comments received from the modelers and the discussions with them concerning the theoretical formulation and performance of their models. We wish to thank:

Mr. Joseph Scire — Environmental Research and Technology
(now at Sigma Research, Inc.)

Mr. Martin Schock — North Dakota Department of Health
Mr. Steven Weber

Dr. I-Tung Wang — Combustion Engineering, Inc.

Dr. Steve Mueller — Tennessee Valley Authority
Dr. Tim Crawford

Mr. Doug Stewart — Systems Applications, Inc.
Mr. Gary Moore
Mr. Ralph Morris

TABLE OF CONTENTS

VOLUME I

ABSTRACT . . . iii
ACKNOWLEDGMENTS . . . v
TABLE OF CONTENTS . . . vii
LIST OF FIGURES . . . xiii
LIST OF TABLES . . . xix

SECTION 1 INTRODUCTION . . . 1-1

- 1.1. BACKGROUND . . . 1-1
- 1.2. MODEL EVALUATION PROTOCOL . . . 1-2
- 1.3. DEFINITION OF PROJECT TASKS . . . 1-3
- 1.4. ORGANIZATION OF THIS REPORT . . . 1-4

SECTION 2 REVIEW OF THEORETICAL FORMULATIONS OF THE MODELS . . . 2-1

- 2.1. INTRODUCTION . . . 2-1
- 2.2. DISCUSSION OF KEY FEATURES OF PREPROCESSOR MODEL THEORIES . . . 2-8
    - 2.2.1. The Meteorological Preprocessor - MESOPAC . . . 2-8
    - 2.2.2. The Meteorological Preprocessor - MSPACK . . . 2-9
    - 2.2.3. The Meteorological Preprocessor - MESOPAC II . . . 2-11
    - 2.2.4. The Meteorological Preprocessor for MTDDIS . . . 2-12
    - 2.2.5. The Meteorological Preprocessor - The BLM/MDPP Model . . . 2-14
    - 2.2.6. The Application of MESOPAC II to the RADM Model . . . 2-16
    - 2.2.7. The Application of MESOPAC and MESOPAC II for RTM-II . . . 2-17
- 2.3. DISCUSSION OF KEY FEATURES OF PLUME MODEL THEORIES . . . 2-17
    - 2.3.1. MESOPUFF Model . . . 2-17
    - 2.3.2. MESOPLUME Model . . . 2-18
    - 2.3.3. MSPUFF Model . . . 2-21
    - 2.3.4. MESOPUFF II Model . . . 2-23
    - 2.3.5. MTDDIS Model . . . 2-25
    - 2.3.6. ARRPA Model . . . 2-26
    - 2.3.7. RADM Model . . . 2-30
    - 2.3.8. RTM-II Model . . . 2-31

SECTION 3 THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA BASES . . . 3-1

- 3.1. INTRODUCTION . . . 3-1
- 3.2. METEOROLOGICAL DATA REQUIREMENTS OF THE MODELS . . . 3-1
- 3.3. OKLAHOMA FIELD EXPERIMENTS . . . 3-3
    - 3.3.1. Description of Experiments . . . 3-3
    - 3.3.2. Ground Level Tracer Observations . . . 3-4
    - 3.3.3. The Oklahoma Modelers' Data Base . . . 3-11

Preceding page blank

TABLE OF CONTENTS (CONTINUED)

3.3.3.1. The Meteorological Grid System 3-11
3.3.3.2. Surface Weather Observations / Modelers' Data Base. 3-14
3.3.3.3. Rawinsonde Observations / Modelers' Data Base. 3-14
3.3.3.4. Meteorological Tower Observations / Modelers' Data Base. 3-16

3.3.4. The Treatment of Missing Data. 3-16
3.3.5. The Source Release / Modelers' Data Base 3-19

3.4. SAVANNAH RIVER PLANT KRYPTON-85 EXPERIMENTS. 3-19

3.4.1. Description of Experiments 3-19
3.4.2. Ground-Level Krypton-85 Observations / Study Periods . . 3-20
3.4.3. The Savannah River Plant Modelers' Data Base 3-20

3.4.3.1. The Meteorological Grid System 3-20
3.4.3.2. Surface Weather Observations / Modelers' Data Base. 3-22
3.4.3.3. Rawinsonde Observations / Modelers' Data Base. 3-24
3.4.3.4. Meteorological Tower Observations / Modelers' Data Base. 3-25
3.4.3.5. The Source Release / Modelers' Data Base . . . 3-26

SECTION 4 OPERATIONAL EVALUATION BASED ON MODEL/DATA COMPARISONS . . . 4-1

4.1. INTRODUCTION 4-1
4.2. APPLICATION OF AMS STATISTICS TO THE OKLAHOMA AND SRP DATA BASES 4-2

4.2.1. Introduction 4-2
4.2.2. Statistical Data Sets for Comparison of Observed Predicted Concentrations 4-3
4.2.3. Peak Concentrations. 4-5
4.2.4. Comparisons of All Concentrations. 4-6
4.2.5. Statistical Analysis of Model Performance. 4-7
4.2.6. Model Performance Results using AMS Statistics 4-12

4.3. GRAPHICAL EVALUATION OF THE EIGHT MODELS WITH THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA. 4-23

4.3.1. Scatter Plots. 4-26
4.3.2. Other Graphical Comparisons. 4-36

4.3.2.1. Oklahoma Data Base 4-36
4.3.2.2. Savannah River Plant Data Base 4-42

TABLE OF CONTENTS (CONTINUED)

SECTION 5 DIAGNOSTIC EVALUATION BASED ON MODEL/DATA COMPARISONS. . . . 5-1

5.1. INTRODUCTION 5-1
5.2. COMPARISON OF PREDICTED CONCENTRATION ISOPLETHS WITH DATA AT OKLAHOMA 5-2
5.3. COMPARISON OF PREDICTED CONCENTRATION ISOPLETHS AT THE SAVANNAH RIVER PLANT 5-19
5.4. PATTERN COMPARISON METHOD OF MODEL EVALUATION (OKLAHOMA ONLY). 5-28
5.5. SUMMARY OF DIAGNOSTIC REVIEW OF MODEL/DATA COMPARISONS 5-34

  5.5.1. MESOPUFF. 5-40
  5.5.2. MESOPLUME. 5-43
  5.5.3. MSPUFF 5-46
  5.5.4. MESOPUFF II. 5-48
  5.5.5. MTDDIS 5-51
  5.5.6. ARRPA. 5-54
  5.5.7. RADM 5-57
  5.5.8. RTM-II 5-60

    5.5.8.1. RTM-II Features Common to Both the Oklahoma and the SRP Data Bases 5-61
    5.5.8.2. RTM-II Features Specific to the Oklahoma Data Base. 5-62
    5.5.8.3. RTM-II Features Specific to the Savannah River Plant Data Base. 5-63

SECTION 6 SUMMARY AND CONCLUSIONS. 6-1

6.1 OVERVIEW 6-1
6.2. GENERAL PERFORMANCE FEATURES OF THE MODELS 6-2
6.3. QUANTITATIVE MEASURES OF MODEL PERFORMANCE 6-3
6.4. SPECIFIC PERFORMANCE FEATURES OF INDIVIDUAL MODELS 6-5

  6.4.1. MESOPUFF 6-5
  6.4.2. MESOPLUME. 6-6
  6.4.3. MSPUFF 6-6
  6.4.4. MESOPUFF II. 6-7
  6.4.5. MTDDIS 6-7
  6.4.6. ARRPA. 6-8
  6.4.7. RADM 6-8
  6.4.8. RTM-II 6-8

6.5. OPTIONS AND PARAMETERS 6-9

REFERENCES. R-1

TABLE OF CONTENTS (CONTINUED)

VOLUME II - APPENDICES

TABLE OF CONTENTS iii
LIST OF FIGURES vii
LIST OF TABLES. xxix

APPENDIX A APCA PAPER/HARTFORD MEETING/OCTOBER 1983. . . . . A-1

APPENDIX B CHOICE OF MODEL-SPECIFIC INPUTS

INTRODUCTION. B-1

PART I: DATA BASE SELECTION AND INPUT PARAMETERS FOR THE MTDDIS MODEL B-1

Oklahoma Tracer Study. B-1
Savannah River Plant Krypton-85 Study. B-2

Source Information B-2
Meteorological Information B-2
Spatial and Temporal Grids B-3
Model-Specific Options B-5

PART II: DATA BASE SELECTION AND INPUT PARAMETERS FOR THE ARRPA MODEL. B-5

Oklahoma Tracer Study. B-5

Source Information B-5
Meteorological Information B-6
Spatial and Temporal Grids B-6
Model-Specific Options B-6

Savannah River Plant Krypton-85 Study. B-7

PART III: DESCRIPTION OF MESOPAC CHANGES AND OPTIONS. B-7

PART IV: DATA BASE SELECTION AND INPUT PARAMETERS FOR THE MESOPAC II WIND FIELD MODEL B-9

Description of MESOPAC II Options. B-10

PART V: DATA BASE SELECTION AND INPUT PARAMETERS FOR THE MESOPUFF, MESOPLUME, MSPUFF, AND MESOPUFF II MODELS . . B-13

Introduction B-13
Computational Considerations B-13
Decision Points. B-14

PART VI: DESCRIPTION OF RTM-II CHANGES AND OPTIONS B-16

TABLE OF CONTENTS (CONTINUED)

PART VII: DESCRIPTION OF RADM CHANGES AND OPTIONS B-17

APPENDIX C DESCRIPTION OF CODE MODIFICATIONS REQUIRED FOR THE EIGHT MODELS

INTRODUCTION. C-1

C.1. DESCRIPTION OF MTDDIS MODIFICATIONS C-1
C.2. DESCRIPTION OF ARRPA MODIFICATIONS. C-3
C.3. DESCRIPTION OF MESOPAC (METEOROLOGICAL PREPROCESSOR) MODIFICATIONS C-6
C.4. DESCRIPTION OF MESOPUFF MODIFICATIONS C-7
C.5. DESCRIPTION OF MESOPLUME MODIFICATIONS. C-9
C.6. DESCRIPTION OF MSPACK (METEOROLGICAL PREPROCESSOR) MODIFICATIONS C-11
C.7. DESCRIPTION OF MSPUFF MODIFICATIONS C-11
C.8. DESCRIPTION OF MESOPAC II (METEOROLOGICAL PREPROCESSOR) MODIFICATIONS C-13
C.9. DESCRIPTION OF MESOPUFF II MODIFICATIONS. C-17
C.10. RTM-II MODIFICATIONS (OKLAHOMA ONLY). C-18
C.11. DESCRIPTION OF RADM MODIFICATIONS C-20

APPENDIX D COMPLETE STATISTICAL COMPARISONS OF THE EIGHT MODELS WITH THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA BASES

INTRODUCTION. D-1

PART I: TABULAR LISTING OF THE PREDICTED AND OBSERVED GROUND-LEVEL CONCENTRATIONS FOR EACH OF THE SRP AND OKLAHOMA DATA SETS D-2

PART II: PRESENTATION OF COMPLETE AMS STATISTICS RESULTS FOR THE OKLAHOMA AND SRP DATA SETS. D-23

APPENDIX E COMPLETE GRAPHICAL COMPARISONS OF THE EIGHT MODELS WITH THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA BASES

INTRODUCTION. E-1

1. LOCATION OF AIR SAMPLERS AND SITE MAP FOR OKLAHOMA EXPERIMENTS E-3

2. EVIDENCE OF LOW-LYING NOCTURNAL JET DURING THE OKLAHOMA STUDY. E-7

3. ISOPLETH PLOTS OF PREDICTED GROUND-LEVEL CONCENTATIONS FOR THE OKLAHOMA EXPERIMENTS. E-18

4. SUMMARY GRAPHICAL PLOTS COMPARING MODEL PREDICTIONS AND FIELD DATA AT OKLAHOMA. E-103

TABLE OF CONTENTS (CONTINUED)

5. LOCATION OF AIR SAMPLERS AND SITE MAP FOR SAVANNAH RIVER EXPERIMENT. E-132

6. SKETCHES OF PREDICTED AND OBSERVED GROUND-LEVEL CONCENTRATIONS FOR TWO SAVANNAH RIVER PLANT EXPERIMENTS SUBCASES 4B AND 6C . E-134

7. SUMMARY GRAPHICAL PLOTS COMPARING MODEL PREDICTIONS AND FIELD DATA AT SAVANNAH RIVER PLANT. E-143

LIST OF FIGURES

Figure Page

2-1 Schematic representation of puff superposition approach 2-2

2-2 Schematic representation of segmented plume approach. 2-2

2-3 Sample meteorological, computational, and sampling grid 2-4

3-1 Location of the sequential air samplers (BATS) and aircraft sampling path at 100 km from the Oklahoma tracer release site . 3-5

3-2 Location of sequential samplers (BATS), LASL samplers, and aircraft sampling flight path at 600 km from the Oklahoma tracer release site. The locations of rawinsonde stations are also shown. 3-5

3-3 Average 45-min perfluorocarbon (PMCH) concentrations along the 100 km arc from the Oklahoma experiment No. 1 (July 8, 1980) . 3-6

3-4 Average 3-hour perfluorocarbon (PMCH) concentration along the 600 km arc for the Oklahoma experiment No. 1 (July 8, 1980). . 3-7

3-5 Average 3-hour PMCH concentrations along the 600 km arc for the period July 9, 0800 GMT to July 11, 2000 GMT ... Oklahoma experiment No. 1 (July 8, 1980) 3-8

3-6 Average 45-min PMCH concentrations along the 100 km arc from the Oklahoma experiment No. 2 (July 11,1980). 3-10

3-7 Location of significant source, meteorological, and receptor sites for the Oklahoma Cases No. 1 and 2. 3-12

3-8 Location of significant source, meteorological, and receptor sites for the Savannah River Plant data base (includes final region choice; inner box is model solution region) 3-23

4-1 Scatter plot of observed and MESOPUFF predicted concentrations at Oklahoma (points paired in space and time) 4-27

4-2 Scatter plot of observed and MESOPLUME predicted concentrations at Oklahoma (points paired in space and time) 4-27

4-3 Scatter plot of observed and MSPUFF predicted concentrations at Oklahoma (points paired in space and time) 4-28

4-4 Scatter plot of observed and MESOPUFF II predicted concentrations at Oklahoma (points paired in space and time) . 4-28

4-5 Scatter plot of observed and ARRPA predicted concentrations at Oklahoma (points paired in space and time) 4-29

LIST OF FIGURES (CONTINUED)

Figure	Page
4-6 Scatter plot of observed and RADM predicted concentrations at Oklahoma (points paired in space and time) 4-29

4-7 Scatter plot of observed and RTM-II predicted concentrations at Oklahoma (points paired in space and time) 4-30

4-8 Scatter plot of observed and MESOPUFF predicted concentrations at Savannah River Plant (points paired in space and time) . . . 4-31

4-9 Scatter plot of observed and MESOPLUME predicted concentrations at Savannah River Plant (points paired in space and time). . . 4-31

4-10 Scatter plot of observed and MSPUFF predicted concentrations at Savannah River Plant (points paired in space and time) 4-32

4-11 Scatter plot of observed and MESOPUFF II predicted concentrations at Savannah River Plant (points paired in space and time) 4-32

4-12 Scatter plot of observed and MTDDIS predicted concentrations at Savannah River Plant (points paired in space and time) 4-33

4-13 Scatter plot of observed and RADM predicted concentrations at Savannah River Plant (points paired in space and time) 4-33

4-14 Scatter plot of observed and RTM-II predicted concentrations at Savannah River Plant (points paired in space and time) 4-34

4-15 Frequency distribution of predicted and observed concentrations at Oklahoma for MESOPLUME based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}. 4-37

4-16 Cumulative frequency distributions of MESOPLUME predictions and observed concentrations at Oklahoma based on points paired in space and time. 4-37

4-17 Frequency distribution of residuals at Oklahoma for MESOPLUME based on points paired in space and time ... residual range: -100 to 100 parts per 10^{15}. 4-38

4-18 Frequency distribution of predicted concentrations at Oklahoma for MESOPUFF based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}. 4-38

4-19 Frequency distribution of predicted and observed concentrations at Oklahoma for MESOPUFF II based on pints paired in space and time ... concentration range: 0 to 100 parts per 10^{15}. 4-39

LIST OF FIGURES (CONTINUED)

Figure Page

4-20 Frequency distribution of predicted and observed concentrations at Oklahoma for RTM-II based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15} 4-39

4-21 Cumulative frequency distributions of RTM-II predictions and observed concentrations at Oklahoma based on points paired in space and time. 4-40

4-22 Frequency distribution of residuals at Oklahoma for RTM-II based on points paired in space and time ... residual range: -100 to 100 parts per 10^{15}. 4-40

4-23 Frequency distribution of predicted and observed concentrations at Savannah River Plant for MESOPLUME based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3 4-43

4-24 Cumulative frequency distributions of MESOPLUME predictions and observed concentrations at Savannah River Plant basd on points paired in space and time 4-43

4-25 Frequency distribution of residuals at Savannah River Plant for MESOPLUME based on points paired in space and time ... residual range: -100 to 100 pCi/m^3 4-44

4-26 Frequency distribution of predicted and observed concentrations at Savannah River Plant for MTDDIS based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3 4-44

4-27 Cumulative frequency distribution of MTDDIS predictions and observed concentrations at Savannah River Plant based on points paired in space and time 4-45

4-28 Frequency distribution of residuals at Savannah River Plant for MTDDIS based on points paired in space and time ... residual range: -100 to 100 pCi/m^3 4-45

4-29 Frequency distribution of predicted and observed concentrations at Savannah River Plant for MESOPUFF II based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3 . . 4-46

4-30 Frequency distribution of predicted and observed concentrations at Savannah River Plant for RTM-II based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3. . . . 4-46

5-1 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions 5-3

LIST OF FIGURES (CONTINUED)

Figure Page

5-2 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) MSPUFF predictions, (right) MESOPUFF II predictions. 5-4

5-3 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) ARRPA predictions, (right) RADM predictions 5-5

5-4 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... RTM-II predictions 5-6

5-5 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions 5-7

5-6 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... (left) MSPUFF predictions, (right) MESOPUFF II predictions. 5-8

5-7 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... (left) ARRPA predictions, (right) RADM predictions 5-9

5-8 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... RTM-II predictions 5-10

5-9 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions 5-11

5-10 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) MSPUFF predictions, (right) MESOPUFF II predictions 5-12

5-11 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) ARRPA predictions, (right) RADM predictions 5-13

5-12 Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... RTM-II predictions 5-14

5-13 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... (top) MESOPUFF predictions, (bottom) MESOPLUME predictions 5-20

LIST OF FIGURES (CONTINUED)

Figure Page

5-14 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... (top) MSPUFF predictions, (bottom) MESOPUFF II predictions 5-21

5-15 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... (top) MTDDIS predictions, (bottom) RADM predictions. 5-22

5-16 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... RTM-II predictions. 5-23

5-17 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... (top) MESOPUFF predictions, (bottom) MESOPLUME predictions 5-24

5-18 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... (top) MSPUFF predictions, (bottom) MESOPUFF II predictions 5-25

5-19 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... (top) MTDDIS predictions, (bottom) RADM predictions. 5-26

5-20 Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... RTM-II predictions . . 5-27

LIST OF TABLES

Table Page

2-1 Comparison of Major Features of Meteorological Preprocessors. . 2-5

2-2 Comparison of Major Features of Plume Models. 2-6

3-1 Summary of Meteorological (and Land Use) Data Required by the Eight Long-Range Transport Models 3-2

3-2 Timing Schedule of Releases, Ground-Level Samples, and Computational Periods for Oklahoma Cases No. 1 and 2. 3-15

3-3 Availability and Periods of Record for the Eight Rawinsonde Locations for Oklahoma Cases No. 1 and 2. 3-17

3-4 Sample Collection Periods and Model Calculational Periods for 10-Hr Samples ... Representing 15 Data Sets for the Savannah River Plant Krypton-85 Data Base. 3-21

4-1 Summary of Data Sets for Use with AMS Statistics (Examples Given for SRP Data Base Only). 4-4

4-2 Statistics Recommended by AMS for Application to Data Sets Representing All-Concentration Comparisons 4-8

4-3 Statistics Recommended by AMS for Application to Data Sets Representing Peak Values Unpaired (25 Highest) 4-9

4-4 Statistical Estimators and Basis for Confidence Limits for General Performance Measures Recommended by AMS 4-10

4-5 Statistical Data Set (A-1) for Oklahoma Data .....(parts per 10^{15} . (Compares highest observed value for each event with highest prediction for same event, paired in time, not location) 4-13

4-6 Statistical Data Set (A-1) for Savannah River Plant Data (pCi/m^3). (Compares highest observed value for each event with highest prediction for same event, paired in time, not location) 4-14

4-7 Statistical Data Set (A-2) for Oklahoma Data (parts per 10^{15}). (Compares highest observed value over all 21 experimental periods at each monitoring station with the highest prediction for those 21 experimental periods at the same station, paired in location, not time). 4-17

Preceding page blank

LIST OF TABLES (CONTINUED)

Table Page

4-8 Statistical Data Set (A-2) for Savannah River Plant Data (pCi/m^3). (Compares highest observed value over all 65 experimental periods at each monitoring station with the highest prediction for those 65 experimental periods at the same station, paired in location, not time) 4-18

4-9 Statistical Data Set (A-4a) for Oklahoma Data .... (parts per 10^{15}) (Compares highest N (=25) observed and highest N predicted values, regardless of time or location.) 4-20

4-10 Statistical Data Set (A-4a) for Savannah River Plant Data (pCi/m^3). (Compares highest N (=25) observed and highest N predicted values, regardless of time or location.) 4-21

4-11 Statistical Data Set (B-3) for Oklahoma Data (parts per 10^{15}) (Compares observed and predicted values at all stations, paired in time and location.) 4-24

4-12 Statistical Data Set (B-3) for Savannah River Plant Data (pCi/m^3) (Compares observed and predicted values at all stations, paired in time and location.) 4-25

4-13 Comparison of Predictions of the Models Based on Percentage of Pairs where Observed Values are Greater than Predicted Values (points paired in space and time) 4-35

4-14 Summary of Results of Graphical Comparisons of Models with Oklahoma Data 4-41

4-15 Summary of Results of Graphical Comparisons of Models with Savannah River Plant Data 4-47

5-1 Frequency Distribution of Predicted and Observed Concentrations for the Two Oklahoma Cases (Based on Predicted/Observed Pairings in Space and Time) in parts per 10^{15} 5-18

5-2 Frequency Distribution of Predicted and Observed Concentrations for the 15 Savannah River Plant Cases (Based on Predicted/Observed Pairings in Space and Time) in pCi/m^3. . . . 5-29

5-3 Pattern Comparison Results for the 100 km Arc ... July 8, 1980; Oklahoma Experiment 5-32

5-4 Pattern Comparison Results for the 600 km Arc ... July 8, 1980; Oklahoma Experiment 5-32

5-5 Pattern Comparison Results for the 100 km Arc ... July 11, 1980; Oklahoma Experiment 5-32

LIST OF TABLES (CONTINUED)

Table Page

5-6 Summary of Peak Predicted Concentrations for Oklahoma Data Sets, July 1980 (parts/10^{15}). 5-36

5-7 Summary of Peak Predicted Concentrations for Savannah River Plant Data Sets, pCi/m^3 5-37

5-8 Classification of Model Prediction Types for the Savannah River Plant 5-39

5-9 Mixing height at plume center (rounded to 100 m) predicted by MESOPAC II for the first release at Oklahoma. 5-62

SECTION 1

INTRODUCTION

1.1. BACKGROUND

The U.S. Environmental Protection Agency (EPA) is currently evaluating the performance of models in several categories for the purpose of updating EPA's "Guideline on Air Quality Models".[1] Long-range transport models are needed in calculations made at distances beyond about 50 km, the limits of applicability of straight-line Gaussian models. Argonne National Laboratory (ANL) and the University of Illinois (UI) have been under contract with EPA to assist in the evaluation of the eight long-range transport models submitted to EPA in response to the March 27, 1980 Federal Register notice.[2] The evaluation was conducted by testing the models against field data with performance defined largely through the use of statistical measures recommended by the American Meteorological Society (AMS).[3] Additional statistical and graphical techniques were used to supplement the AMS statistics in the performance evaluation.

The eight models submitted were all short-term long-range transport models. These models predict short-term averages (3, 12, 24 hr) of SO_2, sulfates, and particulates over distances of 20-2000 km from single or multiple sources. The EPA required that data bases used to test the models represent single sources in which concentrations above background can be detected over mesoscale distances.

Two data bases were chosen as the basis for the evaluation of the models. The first was the Oklahoma data base[4] with two data sets and the second was the Savannah River Plant (SRP) krypton-85 data base[5] with fifteen data sets. The Oklahoma field studies took place in July 1980 whereas the SRP data base was obtained during the period October 1976 - July 1977. Both data bases involved tracer releases; as a result, only the transport and diffusion components of the models could be tested. Evaluation of the treatment of chemical conversion and wet/dry deposition was beyond the scope of this study. The above data bases contained ground-level concentrations measured with averaging times of 45 minutes to 10 hours.

The eight models submitted for evaluation were:

* MESOPUFF[6,7]......... Environmental Research and Technology, Inc. (ERT)
* MESOPLUME[8].......... Environmental Research and Technology, Inc. (ERT)
* MSPUFF[9,10].......... North Dakota Department of Health.
* MESOPUFF II[11-13].... Environmental Research and Technology, Inc. (ERT)
* MTDDIS[14]............ Rockwell International, Inc.
* ARRPA[15,16].......... Tennessee Valley Authority (TVA)
* RADM[17].............. Dames and Moore, Inc.
* RTM-II[18]............ Systems Applications, Inc. (SAI)

These models are listed in the order that their theoretical development is described in Section 2.

Except for ARRPA and MTDDIS, all models were evaluated with both data bases. The evaluation of ARRPA was limited to the Oklahoma data base because this model employs the Boundary Layer Model (BLM) of the National Weather Service (NWS) for its wind field. BLM output is not available for most periods earlier than December 1979. The MTDDIS model was evaluated only with the SRP data base. The model is not applicable to the ground-level releases of the Oklahoma experiments.

1.2. MODEL EVALUATION PROTOCOL

In the model evaluation presented here, the models were run as stated in the user's manuals using data that the models ordinarily expected as input. No modifications to the codes were made that might be interpreted as model improvement. Decisions on model input parameters were made in a way that was consistent with model theory and represented the likely choice of an informed user.

The model evaluation protocol also involved the restriction on the modelers from submitting theoretical improvements to their codes after testing had begun, although the submission of corrections of computer coding errors was permitted. In this way, no model had an advantage over any other model. Within this general constraint, however, each modeler was permitted to provide a "final" updated code for evaluation. Most of the models had undergone changes since their original submission to EPA based on the 1980 Federal Register request. It was of interest to all involved to have the most current version of each model evaluated.

A final aspect of the model evaluation protocol was the opportunity of the modelers to comment on the work as it progressed. The project was organized into three tasks (as described below) and each modeler had the opportunity to comment on the procedure and accomplishments at each stage. Modelers were asked for their approval of the results of each task before the next task was undertaken.

1.3. DEFINITION OF PROJECT TASKS

The project was divided into three tasks of which the present report provides the results of Task III. The three tasks were as follows:

TASK I. Preparation of Input for Eight Long-Range Transport Models

This task involved the acquisition and computerization of the pertinent portions of the Oklahoma and SRP data bases. Software was developed to extract from those data bases the required model inputs in the proper format. The end result of Task I was 119 input data sets. The Task I report[19] described the data bases and the methodology of model testing for review by the modelers and by EPA. Comments by the modelers and by EPA were given in an Addendum to the Task I report[20] along with ANL/UI responses.

TASK II. Preparation of Sample Runs for Each Model

This task involved the preparation, for each model, of one sample run for each of the two data bases. These runs were included as part of a Task II report[21] to EPA for review by the model developers. In this way, the model developers verified that their computer codes were implemented on the ANL Ridge-32 minicomputer correctly and that the codes were being run according to the requirements of the user's manuals. Comments by the modelers and by EPA were given in an Addendum to the Task II report[22] along with ANL/UI responses.

TASK III. Preparation of Model/Data Comparisons and Model Evaluation

This task, represented by completion of this report, provided the results of the evaluation of the eight models with the 17 data cases. A draft of this report was submitted to the modelers and EPA for review. Comments were received from the model developers and from the EPA and were used in preparation of this final report document.

The project, therefore, proceeded in stages with reviews of the progress of the project carried out at each step.

1.4. ORGANIZATION OF THIS REPORT

This report consists of two volumes. Volume I presents a summary of the analysis and results. Volume II contains the supporting Appendices. An overview of the contents of the remainder of Volume I now follows.

Section 2 of this volume summarizes the key theoretical differences among the models. A discussion of the major assumptions employed in each meteorological preprocessor and plume model is made. Section 3 describes the characteristics of the Oklahoma and SRP data bases. Described there are all the source, meteorological, and receptor data that were available. Section 3 also describes the subset of available data that were used in this model evaluation study.

Section 4 presents an operational evaluation of the models. For the operational evaluation, the model/data comparisons are evaluated by statistical and graphical methods. Results for both data bases are discussed simultaneously. The AMS statistics are supplemented with isopleth plots, scatter diagrams, frequency histograms, cumulative frequency plots, and pattern comparison results. Section 5 presents a diagnostic evaluation of the models. This diagnostic evaluation involves tracing back model/data discrepancies to theoretical assumptions in the models. Section 6 presents the summary and conclusions of this evaluation work.

As noted above, Volume II of this report contains the Appendices A-E. Appendices A, B, and C support the first several sections in Volume I. Appendix A (also Ref. 23) reprints an earlier paper describing the model evaluation work done for the MESOPUFF and RTM-II models with the Savannah River Plant data. Appendix B describes on a model-by-model basis the key assumptions and choices that were made and problems faced in generating the input data for each model. Appendix C describes the coding changes that were required for each model to make it operational on our computer and to allow the codes to be run with the Oklahoma and SRP data bases. Appendices D and E contain auxiliary statistical tables and graphs that could not be placed in Volume I due to space limitations.

SECTION 2

REVIEW OF THEORETICAL FORMULATIONS OF THE MODELS

2.1. INTRODUCTION

This section provides a brief theoretical review of the eight candidate models. Each of the eight models consists of a meteorological preprocessor and a plume dispersion model. For most models, the meteorological preprocessor computes the wind field on a two-dimensional array of grid points that cover the area of interest. In general, the horizontal components of the wind are computed on an hourly basis based on input meteorological data from twice-daily rawinsonde soundings and/or hourly surface data within the region modeled. For each hour, the meteorological preprocessor computes a two dimensional array of mixing heights and Pasquill-Gifford-Turner (PGT) stability classes.

The wind field models are generally spatial and temporal interpolation schemes that operate on the meteorological data that are input to them. These interpolation methods are not solutions to the continuity and momentum equations. These interpolation schemes, although accurate only in a gross sense, are designed to be a simple and practical compromise to the solution of regulatory type problems.

The wind speeds, mixing heights, and PGT stability classes are input into a plume dispersion code which computes plume transport and dispersion as a function of time. Plume transport schemes used by the models are summarized as follows:

(a) dispersion of a series of Gaussian puffs as illustrated in Figure 2-1 (MESOPUFF, MSPUFF, MESOPUFF II, and MTDDIS),

(b) dispersion of a series of contiguous plume segments where these contiguous segments generalize the straight-line Gaussian plume, as illustrated in Figure 2-2 (MESOPLUME and ARRPA),

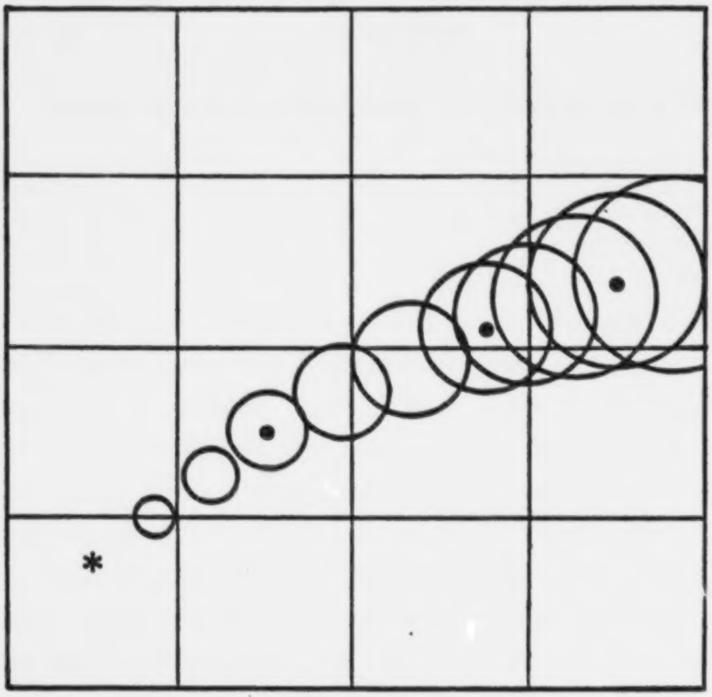


Figure 2-1. Schematic representation of puff superposition approach. (Adapted from Scire, et al.[12]).

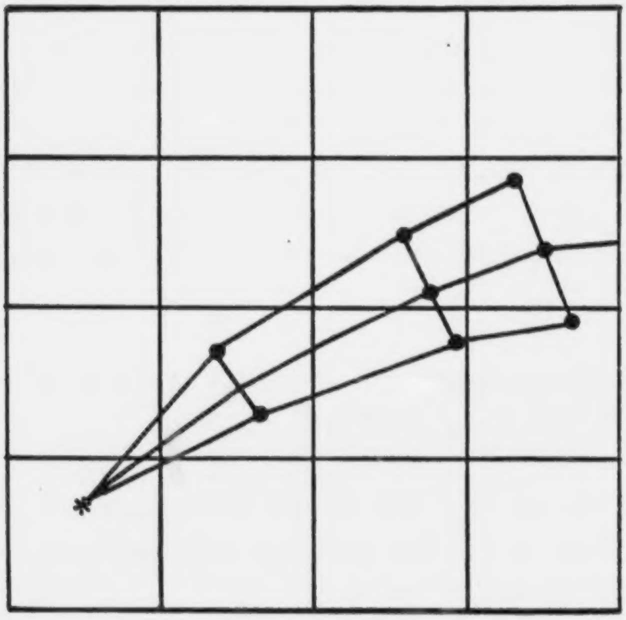


Figure 2-2. Schematic representation of segmented plume approach. (Adapted from Benkley and Bass[8]).

(c) a Lagrangian random walk process in which a large number of particles are released that undergo transport and diffusion (RADM), or

(d) a finite-difference solution to the convective-diffusion equation (RTM-II).

Figure 2-3 shows a sample meteorological grid for which a wind field would be generated. This Figure also illustrates the computational grid for which plume calculations would typically be made. This computational grid is generally a subset of the meteorological grid. A sampling grid is defined as a subset of (sometimes of finer resolution than) the computational grid in which individual gridded and/or non-gridded receptors are located.

Table 2-1 provides a summary of the major features of the meteorological preprocessors employed by the eight candidate long-range transport models. MESOPAC is the meteorological preprocessor for MESOPUFF, MSPACK is used by MSPUFF, and MESOPAC II was developed for MESOPUFF II. The RTM-II model has no fixed meteorological preprocessor and the user must decide which one is most appropriate for the modeling problem at hand. MESOPAC was recommended by the RTM-II model developers for the SRP data base, with MESOPAC II recommended for the Oklahoma data base. For RADM, the model developers recommended the use of MESOPAC II. It will be noted later from Table 3-1 that there is also a wide variety of meteorological data needs among the competing models.

Table 2-2 compares the major features of the plume dispersion models. References 24-41 support Tables 2-1 and 2-2. Described below are the highlights of the theories of the eight candidate models along with the key data requirements of each of the models. The emphasis in the discussion below is in the transport and dispersion portions of the models. The use of tracer releases (Oklahoma and Savannah River Plant data bases) in this evaluation study precludes any testing of the components of the models that deal with chemical transformation and removal processes.

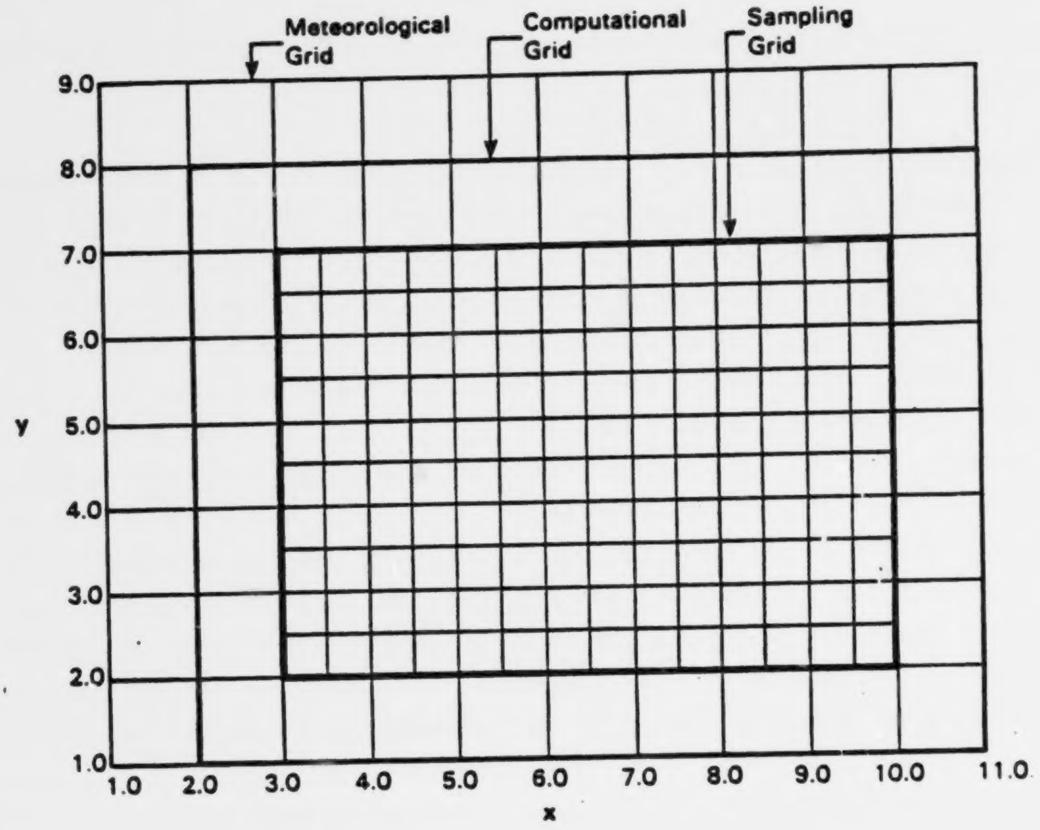


Figure 2-3. Sample meteorological, computational, and sampling grid. (Adapted from Scire, et al. [12])

Table 2-1. Comparison of Major Features of Meteorological Preprocessors

	Input for Wind Field				
Preprocessor	No. of Layers	Layer 1	Layer 2	Mixing Height	Stability Classes
MESOPAC	1	User specified[a]	-	Benkley-Schulman[24]	Modification of Turner[25] using computed mixing height, wind speed (10 m), and time of day at rawinsonde stations
MSPACK	1	User specified[b]	-	Modified Benkley-Schulman[c]	Same as MESOPAC
MESOPAC II	2	Vertically averaged winds (surface to mixing height) or surface winds, or 850 mb, 700 mb, 500 mb winds[d]	Vertically averaged from mixing height to 850 mb, 700 mb, or 500 mb; or single-level wind (e.g. 850 mb, 700 mb, 500 mb)[d]	Maul[26], Venkatram[27,28] specified based on micrometeorological parameters generated from surface data	Turner[29], based on solar radiation, cloud cover data and wind speed at surface stations
MTDDIS	1	Wind speed and wind direction at effective height of plume	-	Interpolation on hourly basis of NDC twice-daily mixing heights based on surface meteorological data	Special algorithm using surface and cloud data including precipitation and time of day providing hourly values of Pasquill categories
ARRPA (BLM/MDPP)	9	Obukhov similarity theory for first layer (0-50 m)	Numerical computations based on surface and upper air data for "transition" layer (50 m-2000 m, sublayers 2-9, levels 2-10)	Modified Benkley-Schulman[e]	Pasquill-Gifford categories computed from Golder[30] diagram using roughness length (z_o) and Monin-Obukhov length (L) output from BLM. The Brookhaven parameterization is computed also based on the vertical potential temperature gradient output from BLM.

[a] For each rawinsonde station and twice-daily period, the user must enter the u and v components of the wind computed by a method of his choice. Possible choices are the single-level winds (850 or 700 mb level) or vertically averaged winds to a fixed height (e.g. 1500 m), etc.

[b] For each rawinsonde station and twice-daily period, the user must enter the u and v components of the wind computed by a method of his choice; however, the user is restricted to single-level winds only (850 mb level is commonly used).

[c] Modifications involve the addition of the treatment of turbulent dissipation of convection in the calculation of the convective mixing height and the use of surface wind data to compute the mechanical mixing height.

[d] The user has the option to choose any combination of the layer 1 and layer 2 alternatives for lower and upper layer winds. Only the most logical choices were presented in the table above.

[e] The Benkley-Schulman method is applied to every 3-hour BLM-predicted wind and temperature profile rather than to the original rawinsonde data.

Table 2-2. Comparison of Major Features of Plume Models

Model	Transport and Dispersion Algorithm		Physical and Chemical Transformation	Removal Processes	
	Modeling Approach	Dispersion Coefficients		Dry	Wet
MESOPUFF	Gaussian Puff[a]	σ_y, σ_z from Turner curves at distances <100 km; σ_y,σ_z from Heffter formulas[f] at distances >100 km	SO_2/SO_4 transformation rate is linear[h], 2%/hr	Deposition velocity concept[h]: v_{d,SO_2} = 1 cm/sec v_{d,SO_4} = 0.1 cm/sec	None
MESOPLUME	Gaussian Plume	σ_y, σ_z from Turner curves at distances <100 km; σ_y, σ_z from Heffter formulas[f] at distances >100 km	SO_2/SO_4 transformation rate is linear[h], 2%/hr	Deposition velocity concept[h]: v_{d,SO_2} = 1 cm/sec v_{d,SO_4} = 0.1 cm/sec	None
MSPUFF	Gaussian Puff[a]	σ_y, σ_z from Turner curves at distances <50 km[b]; Heffter formulas[f] at distances >50 km	SO_2/SO_4 transformation rate is linear[h], 2%/hr	Deposition velocity concept[h,i]: v_{d,SO_2} = 1 cm/sec v_{d,SO_4} = 0.1 cm/sec v_{d,NO_2} = 0.5 cm/sec $v_{d,TSP}$ = 0.5 cm/sec	Wet removal dependent on rainfall rate and scavenging coefficients[h,i]: $\lambda_{SO_2} = 3.0 \times 10^{-5}$ $\lambda_{SO_4} = 1.0 \times 10^{-4}$ $\lambda_{NO_2} = 1.0 \times 10^{-4}$ $\lambda_{TSP} = 1.0 \times 10^{-5}$
MESOPUFF II	Gaussian Puff[a]	σ_y, σ_z from Turner curves at distances < X_o km determined by the user; σ_y, σ_z from Heffter formulas[f] at distances > X_o km (Default value of X_o = 100 km.)	Conversion rates (%/hr) for SO_2 to SO_4, NO_x to HNO_3 + PAN, and NO_x to HNO_3 (only) are computed[e] based on curve fits to hourly (daytime) conversion rates predicted by a photochemical model. Small constant transformation rates are used during nighttime hours	Three-layer model of dry deposition based on deposition velocity concept for SO_2, NO_x, HNO_3, SO_4, and NO_3. Deposition velocities are dependent upon land use and stability class	Wet removal dependent on rainfall rates and scavenging coefficients. Coefficients for SO_2, SO_4, NO_x, HNO_3, NO_3 are dependent on character of precipitation (liquid, frozen)
MTDDIS	Gaussian Puff	σ_y from Heffter[31] formula[f]; σ_z from Draxler[32] formula	SO_2/SO_4 transformation rate is linear, 3%/hr	Deposition velocity concept: v_{d,SO_2} = 1 cm/sec	Semi-empirical wet deposition model that uses observed rain cloud base height and a precipitation rate dependent washout ratio

Table 2-2. Continued

Model	Transport and Dispersion Algorithm		Physical and Chemical Transformation	Removal Processes	
	Modeling Approach	Dispersion Coefficients		Dry	Wet
ARRPA	Gaussian Plume	Composite method[c]	SO_2/SO_4 transformation rate is linear with the rate constant having a diurnal and annual dependency	Deposition velocity concept for SO_2 and sulfate with dependency on stability	None
RADM	Lagrangian Random Walk	K_H is constant with elevation, K_z has a vertical profile formulation[g]	SO_2/SO_4 transformation rate is linear[h]	Deposition velocity concept for SO_2[j]	None
RTM-II	Finite Difference	Horizontal eddy diffusivities by Smagorinsky[33]	SO_2/SO_4 transformation rate is linear with diurnal and latitude dependency	Deposition velocity concept: flux computed through the surface layer with absorption/adsorption at the ground (vegetation dependent)	SO_4 deposition spatially and temporally variable using Scott's[34] parameterization; washout and rainout from three cloud types are simulated. SO_2 deposition spatically and temporally variable using equilibrium parameterization of Hales and Sutter[35]

---

[a] User has option of Gaussian distribution in vertical direction with reflection or instantaneous mixing to full depth of mixed layer.

[b] Turner curves were modified in the range 0-50 km so that MSPUFF would agree with the short-distance MPTER model[36] under steady state conditions.

[c] Horizontal dispersion: Turner curves for $\sigma_y < 1000$ m; for 1000 m $< \sigma_y < 6000$ m, combined Turner curves and Gifford's Lagrangian theory; for $6000 < \sigma_y < 10,000$ m, Gifford's Lagrangian theory[37]; $\sigma_y > 10,000$ m, Taylor's statistical theory.
Vertical dispersion: Turner curves for plume center height < 50 m; Brookhaven curves for plume center height > 50 m.

[d] Emission is released as series of Gaussian puffs. Each puff becomes integrated into the grid system once it grows to the size of a grid cell.

[e] Conversion rates are a function of background ozone concentration, stability class, relative humidity and total solar radiation.

[f] K_z is a function of stability class for MSPUFF, MESOPUFF II, and MTDDIS. K_z is a constant for MESOPUFF and MESOPLUME.

[g] The dispersion coefficients K_H and K_z are taken from McRae et al.[38], Shir[39], O'Brien[40], and Wyngaard[41]. Both K_H and K_z depend on stability through the Monin-Obukhov length. They also depend on surface roughness, surface wind velocity, mixing height, surface friction velocity, the convective velocity scale, and (for stable conditions) the scale height.

[h] Values presented (if any) represent default values. The user defines his choices on model input.

[i] The deposition velocity and scavenging coefficients presented are default values for summer. Winter default values for all eight are one half the summer values.

[j] User supplies reference height and deposition velocity. Parcels below that height incur some mass loss dependent on that height, the deposition velocity assumed, and the time period (1 hour).

2.2. DISCUSSION OF KEY FEATURES OF PREPROCESSOR MODEL THEORIES

2.2.1. The Meteorological Preprocessor - MESOPAC

A meteorological preprocessor to the MESOPUFF and MESOPLUME model called MESOPAC employs twice-daily rawinsonde soundings to provide a gridded field of winds (u,v), mixing heights, and Pasquill-Gifford-Turner (PGT) stability classes on an hour-by-hour basis. MESOPAC provides only a single-layer wind field; i.e. u and v are constant with height but vary in the horizontal plane and with time.

The user's manual to MESOPAC leaves to the user the determination of the most appropriate method of defining the layer for computing the wind field for his application. This determination is crucial to the application of the model. Suggestions are made that the user consider employing the 850 mb winds (if mandatory pressure levels are only considered) or the 700 mb winds (if the terrain is high). Once the user has decided how to define his single-layer wind field, the user must then program that feature into MESOPAC. In this sense, the model is not a simple black box but requires significant engineering judgment in applying the model correctly, with accurate model predictions dependent in good part upon the accuracy of the judgment made.

Winds (and mixing heights) are first determined (twice daily) at each rawinsonde station; from those values, $1/r^2$-interpolation (where r is the distance between a rawinsonde station and a grid point) in space and linear interpolation in time is used to compute the winds (and mixing heights) at each grid point (x,y) at any hour in the day. The mixing depth algorithm is derived from simplifying and generalizing, for regional scales, the site specific algorithm of Benkley and Schulman.[24] At each nodal point on the meteorological grid, the mixing depth is computed as the larger of two independent values: a mechanical value and a convective value. MESOPAC also produces gridded fields of PGT stabilities for each model time step at each grid point. The normal method for computing PGT stability (Turner[25]) relies on wind speed and solar insolation data during the day and wind speed and cloud cover data at night. Wind speed at each grid point is readily available in MESOPAC, but solar insolation and cloud cover are not. Assumptions are made to try to develop realistic stability classes from the data available: (a) wind speed at 10m, (b) computed mixing depth, and (c) the daytime or nightime period. In MESOPAC, then, hourly gridded values are computed for the stability class, mixing height, and the two horizontal wind components.

For the evaluation of MESOPUFF with the Oklahoma and SRP data bases, the model developers recommended that, of the wind field options available, the wind field be computed based on a vertical average of wind speed and direction up to 1500 m height for both the 00Z and 12Z rawinsonde soundings. This assumption was used in MESOPAC for the MESOPUFF and MESOPLUME runs (Oklahoma and SRP sites), and for the RTM-II runs (SRP site only).

The principal data requirements for the MESOPAC model are described below. Winds are predicted using information obtained only from twice-daily rawinsonde data. The soundings used must be for locations inside the meteorological grid being modeled and the data employed from these rawinsondes are wind components, temperatures, and pressures. Due to the way the mixing depth algorithm works, pressure and temperature data are only necessary for the morning (12Z) soundings. MESOPAC is coded to account for missing data. For example, the code will ignore a missing sounding unless all soundings are missing for a 12-hour time step.

2.2.2. The Meteorological Preprocessor - MSPACK

MSPACK is a revision of Version 2.0 of MESOPAC. Version 2.0 of MESOPAC was specially prepared by ERT for a regulatory compliance case in North Dakota.

The major difference between the versions of the two models evaluated here is that MSPACK employs surface wind data in addition to the twice-daily upper air data. However, the surface wind data are not used to predict the wind field, but are used in generating the mechanical mixing heights in the Benkley-Schulman algorithm used by MSPACK.

The key technical changes made to MESOPAC (in developing MSPACK) were:

(a) meteorological stations located outside the meteorological grid may be considered in predicting the wind field. MESOPAC did not permit upper air stations outside the MESOPAC meteorological grid,

(b) the Benkley-Schulman mixing height routine was modified to include turbulent dissipation of convection. Rather than allowing an immediate collapse in the convective mixing height after hour NHRZ (reference hour of afternoon maximum temperature), and the resulting discontinuity in the diurnal mixing height pattern, MSPACK interpolates between the actual convective mixing height at hour NHRZ and a value of zero four hours later. This allows a smoother transition to the mechanical-dominated nocturnal mixing height.

(c) related to (b), the MSPACK program computes the convective mixing depth on the basis of the temperature at hour NHRZ (user input), which is not required to be 00Z (as in MESOPAC). However, NHRZ must be a daytime hour between 13Z and 00Z, inclusive. The model authors commonly use 1600 hours local time for NHRZ for runs in North Dakota,

(d) for each hour, the surface temperature data is saved on disk along with other MSPACK output so that it can be used to compute the buoyancy flux on a hour-by-hour basis in MSPUFF. In MESOPAC, a single value of the buoyancy flux was read for each source and that value remained constant with time,

(e) hourly surface winds have been included in the calculation of mechanical mixing depth; however, the 850 mb (or other user-selected height) wind data are now required in determining the u and v wind fields for plume advection, and

(f) an hourly precipitation field is now produced for each grid point and passed on to the input of MSPUFF. Wet deposition was not treated in MESOPUFF.

The principal data requirements for both MESOPAC and MSPACK are the same except that MSPACK also requires

(a) hourly surface wind speed, wind direction and ambient temperature data, and

(b) precipitation data used for the treatment of wet deposition.

2.2.3. The Meteorological Preprocessor - MESOPAC II

One of the major differences between MESOPAC and MESOPAC II is that MESOPAC II has a two-layer wind field based on both hourly surface data as well as twice-daily rawinsonde data. The lower layer simulates boundary layer flow and the upper layer represents flow above the boundary layer. The user has several options to choose from in defining the winds in either of these two layers. The options are: (1) surface winds (2) vertically averaged winds (surface to mixing height, mixing height to 850 mb, 700 mb, or 500 mb); or (3) winds at mandatory rawinsonde levels (850 mb, 700 mb, or 500 mb). Any combination of choices is available to the user for the lower and upper layers.

The mixed-layer averaged wind and the mixed-layer to 700 mb averaged wind are the current defaults for the lower and upper level winds, respectively. Mixed-layer averaged wind fields are constructed from both hourly surface data and twice-daily rawinsonde data using a scheme adapted from Draxler[32]. In this method, mixed-layer averaged winds at the rawinsonde stations are used to adjust surface winds that have been interpolated to grid point values.

MESOPAC II predicts the mixing layer height as a function of hour in the day from micrometeorological parameters. The daytime (convective) mixing height, the neutral (shear-produced) mixing height, and the stable boundary layer mixing height are computed from such variables as the sensible heat flux, the convective velocity, and the friction velocity. These latter variables are computed from surface data that have been input to the model.

MESOPAC II also has an option to process precipitation data for use in wet deposition calculations in MESOPUFF II.

Three types of principal data input are required for the meteorological preprocessor, MESOPAC II.

The first type are the twice-daily rawinsonde soundings available from the National Climatic Center (NCC). At each sounding level, the following variables are extracted: pressure, height, temperature, wind speed and wind direction.

The second type of data required are the hourly surface meteorological data encompassing (from the CD-144 format): cloud cover, ceiling height, precipitation type, wind speed, wind direction, surface pressure, temperature, and relative humidity. Surface data are required to be input at hourly intervals.

The third type of data allows a classification of the typical surface characteristics in each grid square. Most users will input land use categories rather than having the specific data requested; i.e., roughness lengths and canopy resistances.

2.2.4. The Meteorological Preprocessor for MTDDIS

The meteorological preprocessor for MTDDIS employs only surface wind data. Also, the MTDDIS model is not applicable to surface releases as in the Oklahoma experiment.

The model is unique among the eight evaluated in this report due to its exclusive use of surface wind data in predicting plume transport. The justification given by the model authors for the exclusive use of surface wind data to transport the plume are:

(a) hourly observations of surface winds from a much denser monitoring network are available as opposed to the upper air sounding network. Upper air data are available only at 12-hour intervals from stations that are sparsely distributed, and

(b) for short distances of model application (no more than a few hundreds of kilometers), surface data may be adequate for wind field determination.

Consequently, no twice-daily rawinsonde data are used in the MTDDIS meteorological preprocessor. The preprocessor does, however, require twice-daily mixing heights as available from the National Climatic Center (NCC). The model provides a special interpolation scheme in space and time using the NCC mixing heights (different from that used in the CRSTER Model[42]) for the computation of hourly values of the mixing height.

The single-layer MTDDIS wind field is derived from surface station data by extrapolation upwards of the hourly wind speed and wind direction using wind speed power law exponents and wind directional shear coefficients. Wind direction is assumed to vary linearly with height with the shear coefficient providing the rate of increase (or decrease) above the 10-m surface level. The wind speed power law exponent and the wind directional shear coefficient are determined at each surface station for each hour based on the local stability class. A special algorithm using wind speed, ceiling height, precipitation factors, and time of day is used to provide hourly values of the Pasquill stability class A-F (G stability is treated the same as F). The wind speed power law exponents and shear coefficients are then determined based on stability class from tabular values of these coefficients. A default table is provided in the computer code, but the user may optionally use his own site-specific tabular data.

The speed of the single-layer wind field (at each surface station at each hour) is determined by evaluating the wind-speed power law formula at the effective height of rise of the plume segment released for that hour (one segment per hour). This height of rise equals the stack height plus the buoyant plume rise. Buoyant plume rise is determined from the Briggs formula. The hourly wind direction at each station is taken as the value computed from a linear relationship that includes the 10-m wind direction and the shear coefficient. This formula is evaluated not at the effective height of rise but at the stack height. Interpolation in space and time leads to the development of a two-dimensional wind field that varies in time.

Mixing heights play an important role in the MTDDIS model because they are used to determine the lid used in dispersion calculations. From the NCC mixing heights, a model-specific algorithm is used to derive mixing heights at each station for each hour. The algorithm takes into consideration ceiling height, cloud data, stability, and precipitation in order to perform interpolation between the available morning and afternoon heights. The derived hourly mixing heights are subsequently used in trajectory and dispersion calculations. The MTDDIS treatment of mixing height differs from the CRSTER method largely in its treatment of stable conditions. MTDDIS limits the mixing during all hours of stable conditions rather than only in the hour prior to sunrise.

Since nearly all field data contain some gaps due to missing data, adjustments for missing data are important to the model. MTDDIS provides for a degree of adjustment in such cases. First, incomplete data from hourly surface station records are permitted in order to make use of all available data. Surface station records are flagged as missing for a given hour by assigning a -1.0 value to the wind speed. Such records are subsequently skipped in computations requiring wind speed, and therefore do not count in the trajectory weighting. In some cases, such as for temperature, default values are provided by the code based upon a known reasonable mean value for the study period. Missing mixing heights are patched by selecting the next closest available station with data when making concentration calculations.

The principal data requirements for the meteorological preprocessor to MTDDIS are as follows. The model uses four general types of data for model input. Meteorological data come primarily from National Weather Service surface stations (WBANs). Wind speed, wind direction, temperature, ceiling height, and precipitation are among the variables contained in the surface station data. In addition, the model uses mixing heights available from the National Climatic Center (NCC) from tape input. As an option, the user may also provide analyzed mixing heights to complement the NCC values.

2.2.5. The Meteorological Preprocessor - The BLM/MDPP Model

A unique feature of the ARRPA Model is its reliance on the National Weather Service (NWS) Boundary Layer Model (BLM) for the prediction of the wind field. The BLM is a large-scale, atmospheric boundary-layer prognostic model. Horizontal (u,v) and vertical (w) components of the wind field are predicted on an hourly basis on a fixed 80 km x 80 km grid. The BLM model is currently applicable only to the eastern two-thirds of the United States. Predictions of the BLM model are available on tape only through the National Weather Service Techniques Development Laboratory and the Air Quality Branch, Tennessee Valley Authority.

The meteorological preprocessor to ARRPA is actually a postprocessor to BLM. This interface is a program called MDPP. First, a short summary of the BLM model and its output are presented. The BLM model has 10 computational levels, or nine layers. The lowest layer (0-50 m) is simulated using Obukhov similarity theory. The remaining eight layers (50-2000 m) are simulated numerically using approximately 65 upper air observations, along with a number of surface station reports. Numerical computations for mass, momentum, energy, and moisture are solved in the upper transition layer (50-2000 m) using finite difference methods. Only six BLM output parameters are used to determine the information needed by the ARRPA model. These are u, v, w, θ (potential temperature), u* (friction velocity) and L^{-1} (L is the Monin-Obukhov stability parameter). Roughness lengths for each grid point are maintained in a DATA statement in the ARRPA meteorological preprocessor code, MDPP.

Mixing heights are computed in the meteorological preprocessor using a modified form of the Benkley-Schulman method. The modification involves the application of the method to every 3-hour BLM-predicted wind and temperature profile rather than to the original rawinsonde data.

Pasquill-Gifford categories are used in ARRPA to simulate near-source plume dispersion due to small-scale processes. The stability classes are computed from the Golder[30] diagram using roughness lengths and Monin-Obukhov lengths. These stability classes are used for vertical dispersion (plume center height) less than 50 m. More elevated dispersion in ARRPA (plume center height greater than 50 m) requires the Brookhaven stability classes. The Brookhaven parameterization is computed from vertical potential temperature gradients obtained from the BLM output.

Some additional meteorological preprocessing is required to prepare the input to the ARRPA plume code. ARRPA manipulates the BLM output and carries out three types of interpolation. The purpose of the interpolation is to make the meteorological data available on a finer grid than the 80 km x 80 km grid used by the BLM model. Two of these interpolations involve vertical profiles and one involves horizontal profiles. The first type is logarithmic interpolation, and it applies to the vertical interpolation of the horizontal wind components below 191 meters. Above that level, a linear interpolation is used for the horizontal wind components. The linear interpolation in the vertical direction is also applied at all levels for wind direction, vertical wind component, and temperature. In this way, meteorological data are available on a finer grid than the 80 km x 80 km grid used by the BLM model.

A four-point inverse-square distance weighting scheme is used by the ARRPA code to interpolate winds within a given horizontal plane. The algorithm takes the four adjacent values obtained from the BLM grid points that are nearest neighbors to the local position and applies inverse distance-squared weighting to obtain the local value.

2.2.6. The Application of MESOPAC II to the RADM Model

The treatment of the wind field for RADM is commonly made by means of a meteorological preprocessor developed by Dames and Moore, Inc. The company had three meteorological preprocessors developed in-house, DEPTH, WINDSRF, and WIND3D. One of these models, WIND3D, predicted the wind (providing all three components, u, v, and w) and mixing heights as a function of time and space. However, all three models were not available to outside users since they had not passed through company quality assurance. A very simple wind and mixing height algorithm was available with the model. This version would have used an internally-computed wind velocity profile for the advection of parcels based only on surface meteorological data. Considering the circumstances surrounding the lack of an adequate/available wind field model for RADM plume predictions, Dames and Moore, Inc. representatives indicated that they wished MESOPAC II to be used with their model. The MESOPAC II model was similar to their best in-house model in terms of the interface to the RADM plume model.

As noted above, the MESOPAC II preprocessor was chosen for the current study by Dames and Moore, Inc. MESOPAC II was programmed by ANL to provide the lower-layer (Layer 1) wind field for use in RADM. The Layer 1 field was obtained by averaging the wind from the ground to the mixing height at each grid point (an option in MESOPAC II). This method of obtaining the wind field and mixing heights for input to RADM was consistent with the usual way Dames and Moore, Inc. applies one of their wind field models in that the modified surface wind includes the true surface wind data, suitably adjusted by the upper air data. It should be recognized that the actual surface wind data (unaltered by upper air data) were still used directly to develop Pasquill stability classes in RADM. As a result, two wind fields were output from MESOPAC II for RADM input; the modified surface winds for plume transport and the true surface wind speed data for Pasquill stability class determination. Mixing heights and surface roughnesses were also output from MESOPAC II for input to RADM. Further details concerning the interfacing of the MESOPAC II meteorological preprocessor with the RADM plume model are discussed in Appendix C.

2.2.7. The Application of MESOPAC and MESOPAC II for RTM-II

There is no single wind-field and mixing height model recommended by the developers, Systems Applications Inc. (SAI), for the RTM-II model. Several wind field models have been used by them in previous model studies. The choice of wind field is dependent upon the nature of the modeling problem. For the Savannah River Plant data base, SAI suggested the use of MESOPAC. Those computer runs were carried out at ANL in 1983. For the more recent application to the Oklahoma experiments (in 1985), MESOPAC II was now available in the literature and was the choice of SAI for that data base. Both these meteorological preprocessors are discussed above. The computer codes for MESOPAC and MESOPAC II were both modified to provide the binary input files on an hourly basis as required by RTM-II.

2.3. DISCUSSION OF KEY FEATURES OF PLUME MODEL THEORIES

2.3.1. MESOPUFF Model

MESOPUFF is a variable trajectory Gaussian puff superposition model. In MESOPUFF, the emission is divided into a series of puffs emitted continually over time. The Gaussian puffs are advected by the wind and dispersed using σ_y and σ_z values from Turner[25] up to 100 km, and from Heffter[31] for distances beyond 100 km. At each hour, the contributions of each puff to the concentration at each sampler is determined. MESOPUFF treats dry deposition and chemical conversion yet does not treat wet deposition.

A Gaussian distribution is assumed for each puff upon exit from the source. MESOPUFF permits the user to specify one of two possible algorithms for the vertical dispersion: (a) a uniform vertical distribution below the mixing height, or (b) a Gaussian, multiple reflection algorithm considering reflection from the ground and the mixing height. Puffs emitted below the local hourly mixing height will contribute to ground-level concentrations. Puffs emitted above the mixing height will have no ground impacts until the height of the mixing layer rises above the puff center. In MESOPUFF, the appropriate mixing height to use in the dispersion algorithm for any puff is the highest mixing height since the release of that puff, in spite of the fact that the mixing height might have been lowered since that particular hour.

The computational scheme of the MESOPUFF model has three distinct functional elements: (1) a Lagrangian puff trajectory function, (2) a puff dispersion function, and (3) a puff sampling function. The Lagrangian trajectory function is used to advect the centerpoint of each puff during a basic time step. The radius of each puff is determined by the puff dispersion function. Given the size and location of every puff, the puff sampling function computes the concentration exposure received at each grid point. This calculation is done for every time step and grid point by summing up the individual puff contributions at that grid point.

Key input requirements for MESOPUFF are the buoyancy flux, the stack height, and the pollutant emission rates for each source in addition to MESOPAC predictions of wind components, mixing heights and stabilities. Use of the buoyancy flux as input does not allow for changing buoyancy with time resulting from emission or ambient temperature changes on an hour-by-hour basis. There are available, however, 24 multiplicative factors that allow for hourly changes in buoyancy flux on a daily basis. In order to run MESOPUFF, a puff release rate and a puff sampling rate must be chosen. Described in the user's manual is an algorithm for the choice of these two parameters based upon wind field predictions from MESOPAC.

2.3.2. MESOPLUME Model

The MESOPLUME Model is a regional-scale variable trajectory Gaussian plume segment model. The MESOPLUME model is very similar to the MESOPUFF model. First, both employ the same meteorological preprocessor, MESOPAC. In addition, the MESOPUFF and MESOPLUME dispersion models are structured similarly. However, the MESOPUFF model treats the plume as a series of superimposed Gaussian puffs whereas the MESOPLUME model treats the plume as divided into contiguous segments. MESOPLUME is, therefore, a generalization of the conventional straight-line Gaussian plume model to regional applications. Each segment of the plume in MESOPLUME describes a portion of the plume between successive time periods; the end points of each segment are advected in a Lagrangian sense. The representation of a continuous plume by the segmented-plume approach is depicted in Figure 2-2. As with MESOPUFF, the MESOPLUME model treats linear conversion of sulfur dioxide (SO_2) to sulfate (SO_4), and dry deposition of SO_2 and SO_4.

In MESOPLUME, a continuous plume is simulated by subdividing the plume into a number of contiguous "plume segment" elements. A simple equation of conservation of pollutant mass for each plume segment helps define the concentration of that plume segment in its transport from the source. The gain or loss by chemical conversion and the loss by dry deposition are treated in that mass conservation equation. The user has the option of specifying two possible vertical distributions for the plume segment: (1) a vertical Gaussian profile, ignoring any effects of the mixing lid, or (2) a uniform vertical distribution below the mixing lid. In Case 1, reflection from the ground is assumed. In Case 2, no ground-level concentrations are calculated if the plume centerline lies above the mixing lid. The relevant mixing height at any time, for a given segment, is the maximum mixing height that the segment has encountered during its travel time from the source.

MESOPLUME requires the following fields (usually at hourly intervals) at each grid point (from MESOPAC):

- horizontal (u,v) wind components
- mixing depth, and
- Pasquill-Gifford-Turner (PGT) stability classes.

The computational scheme of the MESOPLUME model has three distinct functional elements:

(a) a Lagrangian trajectory function...This function is used to advect the endpoints of each plume segment during a basic time step. The resultant distance between consecutive endpoints defines the length of that plume segment,

(b) a plume dispersion function...The widths at the upwind and downwind ends of each plume segment are determined by this function. The plume dispersion parameters, σ_y and σ_z, are calculated for distances less than 100 km using plume growth formulas fitted to the curves of Turner.[24] For distances greater than 100 km, the plume growth rates given by Heffter[31] are used, and

(c) plume sampling function...Given the size and location of each plume segment, the plume sampling function computes the concentration exposure received during that time step for each grid point that lies under the plume segment.

The treatment of the mixing layer is the same as for MESOPUFF with respect to mixing height interaction with the plume segments (rather than puffs).

The MESOPLUME model has three potential problem areas that require attention. First, in the current version of MESOPLUME, no reflection from the mixing height is assumed in Case 1 listed above. Consequently, in Cases 1 and 2, MESOPLUME will not give correct results for distances where the plume has not yet approached a uniform vertical distribution. This problem has the greatest effect under stable conditions because the plume will not approach a uniform vertical distribution until many kilometers downwind. The model developers recommend that the Case 1 option only be used in MESOPLUME at and beyond distances for which the assumption of uniform mixing is appropriate - often at distances of 100 km or more. It is actually a simple modification to incorporate an optional Gaussian reflected vertical distribution so as to make MESOPLUME suitable for near-field computations. (The MESOPUFF model had already been augmented by the model authors to handle reflection from the mixing lid in order to provide meaningful near-field impacts.)

The second area relates to the shrinking of the segment size when the wind flow along the plume axis rapidly decelerates. Due to the inverse wind speed dependence in the concentration algorithm, the predicted concentration for ground-level points beneath the elevated segment can actually increase. This problem is more fundamental and arises from the basic advective-diffusive scheme of the model.

The third area relates to the way in which adjacent plume segments are actually juxtaposed in a curvilinear flow. In the presence of strongly sheared flows (i.e., recirculating flows), adjacent segments may not be perfectly contiguous. Some portions of the segments will overlap causing grid points below to receive two concentration doses during one time step. Other grid points may be passed over without impact at all due to the break of continuity between adjacent segments. This potential problem is also one that is fundamental to the Gaussian segment approach.

The key input requirements to MESOPLUME are the same as for MESOPUFF except that a puff release rate and the puff sampling rate in MESOPUFF are no longer needed. The MESOPLUME user must, however, define a basic time step which determines the plume segments. The MESOPLUME model releases plume segments as a function of time and, as with the release of puffs in MESOPUFF, each segment will eventually pass outside the computational grid and will no longer be considered in the computation.

2.3.3. MSPUFF Model

The MSPUFF Model was developed by the North Dakota State Department of Health (NDSDH) as a proposed improvement of an earlier version of the MESOPUFF Model. The MSPACK/MSPUFF modeling system encompasses a revised MESOPAC (MSPACK) and a revision of MESOPUFF (MSPUFF). The MESOPUFF model evaluated in this report is Version 6.0. MSPUFF is a modification of Version 2.1 of MESOPUFF. (Version 2.0 of MESOPAC and Version 2.1 of MESOPUFF were specially prepared by ERT for a regulatory compliance case in North Dakota.) The version of MSPUFF used by Argonne/UI includes all revisions of MESOPUFF from Version 2.0 through Version 6.0. The MSPUFF Model is commonly used in the State of North Dakota for licensing situations involving long-range transport to Class I areas. The MSPUFF Model has nearly identical input requirements to MESOPUFF as described above and very similar theoretical underpinnings. Major differences between MESOPUFF and MSPUFF are described below.

The key changes to the MESOPUFF Model (in the development of MSPUFF) are:

(a) the changeover in the functional form (Turner curves to Heffter formulas) for σ_y and σ_z occurs at 50 km in MSPUFF as opposed to 100 km as used in MESOPUFF and MESOPLUME. Furthermore, the value of the vertical diffusivity in the Heffter form of σ_z is varied for each stability class. Previously, a constant value of 5 m^2/sec was used in MESOPUFF and MESOPLUME,

(b) the a and b parameters in the power law form of σ_y and σ_z have been adjusted for downwind distances less than 50 kilometers. The revised power law forms now closely agree in the range 2 to 50 kilometers with the values from the steady-state MPTER model.[36] This revision was done to assure agreement between the two models often used in regulatory applications by the North Dakota Department of Health. In North Dakota, the analysis of air quality for short-range distances (less than 50 km) is often made with steady-state models, such as MPTER. Mesoscale range predictions (between 50 km and 250 km) are commonly made with MSPUFF. Extensive model testing has not been carried out to compare MSPUFF and MESOPUFF predictions due to this change in σ_y, σ_z values for MSPUFF. However, in one case comparison (release from a short stack), the MSPUFF predictions were greater than the MESOPUFF predictions for distances less than about 10 km, and lower than MESOPUFF predictions for further distances, and

(c) for those puffs released above the mixed layer, the dispersion has been constrained to the use of dispersion coefficients for stability classes E or F until the mixed layer depth becomes greater than the puff height,

(d) the dependence of dry deposition velocity on stability has been added to the model,

(e) the buoyancy flux, F, for plume rise is computed hourly for each source from the surface temperature at that meteorological station nearest the source. In MESOPUFF, a single buoyancy flux value, F, was read in for each source and was assumed to be applicable to the entire period of model predictions, and

(f) the treatment of sulfur dioxide wet deposition has been added to the model; the washout coefficient for sulfur dioxide is a function of the rainfall rate.

As noted earlier, the input requirements for MSPUFF are nearly identical to MESOPUFF. For MSPUFF (as compared with MESOPUFF), hourly stack dynamic operating parameters and hourly ambient temperatures are employed in predicting source buoyancy flux as a function of hour using the surface station nearest the source.

The model can handle up to four pollutants simultaneously: SO_2, sulfates, NO_x, and particulates. Hourly source emissions data are required for each pollutant (as compared with only SO_2 and sulfates for the MESOPUFF model). All particulates are assumed to be fine particles and are treated as gases.

2.3.4. MESOPUFF II Model

The MESOPUFF II model was developed to enhance the capabilities and flexibilities of the MESOPUFF model including the prediction of secondary aerosols. Extensive modifications had been made to improve the treatment of advection, vertical dispersion, transformation and removal processes.

The MESOPUFF II dispersion model is a Gaussian variable-trajectory puff superposition model that was designed to account for the spatial and temporal variations in advection, diffusion, transformation, and removal mechanisms on regional scales. A continuous plume is simulated as a series of discrete puffs. Each puff is subject to space and time-varying wet removal, dry deposition, and chemical transformation.

In MESOPUFF II, the basic equation of dispersion of a puff is the Gaussian distribution with reflection from the ground and the mixing height. As with MESOPUFF, the dispersion parameters σ_y and σ_z are calculated for puff travel distances less than a user-input distance X_o (100 km default value) with plume growth functions fitted to the curves of Turner.[25] The time-dependent puff growth equation used for distances greater than X_o are those given by Heffter.[31] In the latter equations, the vertical diffusivity, K_z, is a function of stability class.

MESOPUFF II allows three options for determining growth rates for puffs above the boundary layer: (1) E stability rates, (2) F stability rates, or (3) boundary layer stability rates. The default instructions are to use the E stability growth curves for puffs above the boundary layer.

In terms of chemical transformations, a parameterization of the conversion of SO_2 to SO_4 and NO_x to NO_3 is used. These parameterizations include the chemical equilibrium of the $HNO_3/NH_3/NH_4NO_3$ system. Dry deposition is included by means of resistance modeling including options for source or surface depletion. Time and space-varying wet removal is included based on the precipitation rate (liquid or frozen) with a scavenging coefficient based on precipitation type and intensity.

A change from MESOPUFF was in the treatment of vertical dispersion as it relates to dry deposition. Once the plume has become vertically mixed in the far field, the user has an option to employ a three layer treatment of dry deposition. In this option (as a consequence of vertical mixing), σ_z is no longer applicable and stability classes are no longer used; parameterization of vertical dispersion is made in terms of micrometeorological variables with inclusion of a shallow surface layer in order to predict dry deposition from the plume.

A change involves the treatment of spatially and temporally variable rates of chemical transformation, dry deposition, and wet removal. MESOPUFF simulated the transport and chemical conversion of only SO_2 and sulfates and treated only dry deposition. MESOPUFF II treats up to five pollutants: SO_2, SO_4, NO_x, HNO_3, and NO_3 and can predict wet deposition as well in a spatially and temporally variant manner.

A change involves improvement in the treatment of the puff sampling function. In MESOPUFF, there were some problems of poor resolution of puffs in the near field if the puff sampling rate were not specified high enough. The revised sampling function tends to reduce the error in the near field and makes it less likely that the user obtains poor results with the model if incorrect choices are made of the puff release rate and puff sampling rates.

The key data required are emissions from point (and/or area) sources. For each point source are required: location, stack height, stack diameter, exit velocity, stack gas exit temperature, and the emission rate of each pollutant.

One potential problem for the user is that missing data are not treated within the code. The user is required to input replacement values for all missing data. This can be a time-consuming process for a large modeling region with significant amounts of missing surface data.

2.3.5. MTDDIS Model

The MTDDIS Model (Mesoscale Transport and Dispersion Model for Industrial Plumes) is a variable-trajectory Gaussian puff model. The model is an adaptation of the ARL trajectory model by Heffter.[43] MTDDIS was developed to predict short-term impacts from elevated industrial releases over distances that range from tens to a only a few hundreds of kilometers. It does not apply to surface releases as a result of its special treatment of the single-layer wind field. The model also has an option to treat dry and wet deposition.

MTDDIS uses the common elevated Gaussian puff algorithm to predict concentrations. Reflections due to the presence of the ground and mixed layer are assumed. The Lagrangian evolution of trajectories is achieved by tracking the endpoints of each plume segment. Segments are further sub-divided into a variable number of puffs when better resolution and continuity in the concentration field are required. Horizontal dispersion grows linearly with time as in the ARL Heffter model. Vertical dispersion is treated as a function of stability class and time following the procedure of Draxler.[44]

The modeled region is the distance between the ground and the local mixing height. Ground-level concentrations are computed by summing puff contributions at each grid point for each hour. Vertical mixing is treated as uniform beyond downwind distances where σ_z exceeds 1.6 times the mixing height. When the effective height exceeds the local lid, the contribution to the ground-level concentration field is counted as zero. This situation would occur, for instance, if the stack height exceeded the mixing height.

Trajectories are generated by initiating one new trajectory from the source at the beginning of each hour. From that point, the trajectory segment endpoints are updated once each hour. A composite trajectory segment for each hour is obtained by station weighting as adapted from the ARL-Heffter Model.[42] Each trajectory segment is first propagated independently based upon data from each individual surface station that has data available for the hour. The cumulative segment motion is derived through a weighted averaging of the independent motions. The weighting algorithm accounts for distance using inverse square weighting and for directional alignment using angular weighting. Recall that the hourly wind speed and direction used in the trajectory computations are based upon the predicted values of the wind speed and wind direction at the effective plume height and stack height, respectively, and not simply on the 10-m surface values.

The principal data requirements for the MTDDIS model are as follows. Source-related input data include stack height and diameter, updraft velocity, exit temperature, and hourly-averaged emission rate for each hour. Also, the mean sea-level elevation at each station including source latitude and longitude are required. It should be noted that MTDDIS also provides for input of certain deposition-related parameters, but the SRP data base using passive tracers did not employ the deposition modules.

Inputs relating to the modeling domain include surface elevations, roughness heights, boundaries for wind fields, and boundaries for plume predictions.

Finally, inputs are required to prescribe user options for the predictions on a grid including grid origin, grid boundaries, grid spacing, and locations of specially placed samplers. These inputs are options chosen by the user to permit predictions for the desired time-averaging periods and geographical area.

2.3.6. ARRPA Model

The ARRPA Model (Air Resources Regional Pollution Assessment Model) is a single-source segmented-plume model designed to compute pollutant transport, dispersion, and deposition over the regional and interregional scale. It predicts dry deposition and chemical transformation processes for SO_2 and sulfate.

ARRPA uses a segmented Gaussian plume approach to predict airborne concentrations. The Lagrangian evolution of trajectories is achieved by tracking the segment endpoints. A new segment is emitted from the source during each hour. The endpoints of the trajectory segments are updated once each hour.

Terrain topography is inherited from the BLM model. ARRPA uses $1/r^2$ interpolation between BLM grid points in order to obtain terrain elevations at individual receptors. Because of the BLM mesh spacing, ARRPA cannot adequately resolve complex terrain features. However, by specifying receptor heights at particular locations, concentrations are adjusted to simulate terrain effects to some degree.

The divergence adjustments resulting from dynamic features within the BLM model are inherited by ARRPA, and therefore adjustments for shear are necessary. ARRPA computes the vertical wind field deformation (change in w velocity with elevation) over the plume top and bottom edges to arrive at an adjusted σ_z. The horizontal divergence is adjusted by varying the centerline concentration subject to the constraint that it cannot be increased by shear effects only (that would violate entropy).

Plume rise is predicted by classical one-dimensional entrainment theory following the method of Slawson et al.[45] Vertical transport of segment endpoints uses the interpolated wind at the plume centerline. The ARRPA plume model provides two methods as user options for computing horizontal plume transport. Horizontal components may be computed by averaging the horizontal transport computed over the 2σ vertical region about the plume centerline. The other option involves taking the level nearest the centerline to derive a computationally-faster horizontal transport approximation.

Each plume segment is treated as a continuous Gaussian plume. Horizontal dispersion is simulated using an algorithm which has four growth regimes. Turner's curves[25] are used to model the rapid growth regime; i.e. distances downwind to where σ_y becomes greater than 1000 m. When the size of σ_y exceeds 1000 meters at the start of a plume step, a transition is made modeling σ_y by means of a combination of the Turner curves and Gifford's Lagrangian theory[37]. When the starting σ_y for a plume step is between 6000 and 10000 meters, Gifford's Lagrangian theory is used. In the slow growth region beyond 10000 meters, Taylor's statistical theory is used.

The vertical sigmas are computed at low levels (plume center height less than 50 meters) using Turner's curves. For elevated regimes, Smith's Brookhaven curves[46] are used.

The ARRPA plume model limits diffusion between the ground and an upper atmospheric layer according to set of six rules. These rules limit mixing to an upper level of 2000 meters and selects the vertical wind level to use for computing the transport of a plume segment during a time step; such rules are required because the height of the plume segment varies during the step and could possibly rise above the upper lid.

Although ARRPA advances segment endpoints once per hour in order to improve computational efficiency in transport calculations, an internal stepping routine is used to obtain more refined temporal adjustments for ground-level plume predictions. Concentrations are computed from a number of discrete segment positions based on the size of the horizontal sigma as compared to the amount transported perpendicular to the centerline. When the predominant plume transport is parallel to the centerline, the concentration can be computed for the one-hour interval at a given receptor as a simple straight-line Gaussian plume since the relative motion allows use of a single midpoint for time-averaging. However, perpendicular transport must be treated with smaller time steps since the relative motion is two-dimensional; in this case, the one-hour period is divided into steps with plume parameters (such as σ_z and σ_y) varying linearly between steps. Contributions from each step are summed using multiple reflection terms following the method recommended by Turner.[25] After the plume becomes completely vertically mixed, the simpler conventional horizontal Gaussian (vertically uniform) plume formula is used.

The other segmented Gaussian plume model in this evaluation study is the MESOPLUME model. As discussed in Section 2.3.2., the theoretical approach of a Gaussian plume segment model has two potential problems that should be treated. The first relates to the shrinking of the segment size when the wind flow along the plume axis rapidly decelerates. Due to the inverse wind-speed dependence in the concentration algorithm, the predicted concentration for ground-level points beneath the elevated segment can increase. ARRPA addresses the problem caused by the shrinking segment by means of two mechanisms: (a) it expands the horizontal or vertical width during these periods of decelerated winds to prevent increasing concentrations, and (b) it uses the largest segment length in the segment's history for concentration calculations and not the actual reduced segment length.

The other problem relates to the way in which adjacent plume segments are actually juxtaposed in a curvilinear flow. Adjacent segments that overlap can lead to grid points below those segments receiving double doses during a time step. ARRPA treats that problem by first checking to determine whether the angle between adjacent segments is greater than 90 degrees. (It is 180 degrees for a constant-direction flow.) If greater than 90 degrees, the overlapping portion of the more downwind segment of the two is not included in concentration calculations when overlap occurs. If the angle of adjacent segments is less than 90 degrees, the contributions from both segments are included since this case is thought to represent a physical situation (e.g. sudden wind shift from north to south) causing overlapping puffs. This latter case is not considered a modeling problem.

One difficulty that a user should be aware of with ARRPA is in the assignment of mass to the segments. In ARRPA, at the start of each hour, mass is assigned to the leading edge of a new segment based on the release rate for that hour. Segment masses are determined based on masses assigned at the upwind (hour N) and downwind sides (hour N+1) of the segment. A problem arises, however, at the leading and trailing edges of the continuous plume release. The leading segment, for example, will have a zero assigned mass at its upwind edge and a nonzero assigned mass at its downwind edge. Consequently, nonzero concentrations can result from the leading segment (and final segment as well) where zero contributions would otherwise be expected. In that sense, ARRPA is not quite mass conserving. The problem can be eliminated to large degree by adjusting the emission rate at the beginning and end of the continuous release to insure low emission rates at the very start and very end of the period.

The model uses three general types of data as input information. The meteorological input data to the ARRPA plume model is a subset of the output from the National Weather Service Boundary Layer Model. The BLM grid is 3-dimensional (35 by 30 horizontal, 10 vertical) and is defined by a polar stereographic projection. Only six BLM output parameters are used to determine the information needed by the ARRPA model (for each hour): these are u, v, w, θ, u* and L^{-1}. θ is the potential temperature, u* is the friction velocity, and L^{-1} is the inverse of the Monin-Obukhov length. The latter two variables are only stored once per horizontal grid point since they are surface layer values. Since the BLM output contains all meteorological data required by the ARRPA plume model, the effort in running the ARRPA model is reduced to the simple task of selecting the grid section and time period that are required for the particular application.

Source-related input data include stack height, diameter, updraft velocity, exit temperature, and hourly-averaged emission rate for each hour. Also, the mean sea level (MSL) elevation at each station including source latitude and longitude are required. ARRPA also provides for the input of certain deposition-related parameters, but these parameters were not relevant to the current model evaluation program since data from tracer releases were used.

As with the other plume models, inputs involving the characteristics of the study area include receptor elevations, lateral boundaries for wind field model input, and the boundaries for plume model prediction. The ARRPA plume model contains an internal storage of the mean sea level (MSL) elevations for the BLM grid points. Inputs are required to prescribe user options for computation and prediction grid characteristics including grid origin, grid boundaries, grid spacing, and locations of specially placed samplers. The model allows the user to choose a rectangular grid with arbitrary spacing and directional orientation.

2.3.7. RADM Model

The RADM model employs a Lagrangian random walk approach to dispersion. In this method, the mass release is divided up into a number of parcels and these parcels are advected and transported by the wind. Puff advection is treated as in the Gaussian puff models, yet the diffusion employs a displacement by means of a random walk that is scaled by horizontal and vertical dispersion coefficients.

The random walk procedure transports and disperses parcels after release into the meteorological grid region. The horizontal and vertical eddy diffusivities are used to scale the random contribution to each parcel's motion, whereas the layer-averaged winds are used to provide the advective contribution to each parcel's motion. The horizontal eddy diffusivity is calculated at each grid point, but is not a function of height. The vertical eddy diffusivity is also gridded, and is determined for each of up to 8 layers. In the calculations for the Oklahoma and Savannah River Plant data sets, six layers were used for diffusivity computation; each layer requires a complete grid of diffusivity values to be stored. Additional layers would not add important detail. Parcel contributions to pollutant concentrations are summed at the end of each hour for each grid point based on the fractional overlap of the receptor volume with an equal volume centered on each parcel. In the transport of parcel mass, RADM assumes a simple model of chemical conversion and of dry deposition as it follows the dispersion of two pollutants (typically SO_2 and SO_4). A user-supplied linear decay rate (independent of space and time) is used to simulate chemical conversion of pollutant 1 to pollutant 2 and for the decay of pollutant 2. For parcels sufficiently near the ground, a user-supplied deposition velocity independent of space and time is used to subtract deposited mass fractions from the parcels. No wet deposition is simulated in the model. In the current evaluation, neither the decay nor deposition portions of the model were tested since passive tracers were used in the field experiments.

Since the model employs the MESOPAC II meteorological preprocessor, the input requirements for RADM will include those for MESOPAC II. A discussion of the input requirements and the modeling methodology for MESOPAC II are presented in Section 2.2.3. The RADM plume model requires a one-time entry of gridded surface roughness heights determined from land use categories as in MESOPAC II.

The RADM plume model requires an average value per hour (determined from available surface station data) of the surface temperature, the total opaque sky cover, the cloud ceiling height, the potential temperature lapse rate in the mixed layer and the potential temperature lapse rate in the inversion. For the latter two variables, the model developers suggested that the values of 0.0 and 0.005 deg C/100 meters may be used. Both variables enter the Briggs plume rise formulas; the effects of the choice of those two parameters was insignificant for the Savannah River Plant case runs and were not used for the ground-level release at Oklahoma. Hourly emissions data with appropriate stack and ambient conditions must also be supplied. Finally, a table of 10000 properly selected random numbers must be read in for the random walk calculation.

2.3.8. RTM-II Model

The RTM-II model uses a puff-on-grid hybrid technique. The emission is released as a series of puffs and these puffs are transported and dispersed using a Gaussian puff formulation. When a puff expands to the size of either dimension of a grid cell (W_x, W_y or W_z = mixing height), the puff is released into the grid system and then disperses according to the finite-difference solution to the convective-diffusion equation. The numerical scheme used is from Boris and Book.[47] This method is characterized by a two-stage, flux corrected transport algorithm in which a correction of pollutant fluxes in the second stage counteracts the numerical diffusion arising from the first stage. Horizontal eddy diffusivities are parameterized in the model using a formula proposed by Smagorinsky[33], where the horizontal diffusivities are equal to an empirical constant (η, based on scale considerations) multiplied by the magnitude of the velocity deformation tensor. In this formulation, $K_H = K_x = K_y = \eta|\mathrm{Def}|$. Furthermore, the horizontal diffusivities are constrained by maximum and minimum values chosen from regional-scale field studies.

The wind field in the model is represented by a single-layer; however, an upper layer is retained to hold plume mass that escapes the mixing layer when the mixing height decreases. Later, when the mixing height increases again, either due to diurnal variation or to advection into a region with a larger mixing height, this plume mass can re-enter the mixing layer. However, transport in this upper layer proceeds according to the same wind field as is used in the lower layer. Although the wind field is single-layer, the model makes some attempt to include advection of air in the vertical when the single-layer winds show either divergence or convergence. At a grid point where the wind field has a net divergence, the air mass needed to conserve mass is considered advected down from the upper layer, bringing part of the upper-layer concentration with it. At a point where the wind field has a net convergence, some mixing layer concentration is removed into the upper layer, representing an upward advection of air. Since MESOPAC II and MESOPAC both predict wind fields without adjusting them to be divergence-free, this theoretical feature involving conservation of mass plays a role in the predictions reported in this study. That feature may represent important physics that could have an impact on model predictions, especially when fronts approach the area.

The convective-diffusion equation involves concentration explicitly, so no special procedure is needed to derive concentration values. The basics of the method were described in the introduction to the RTM-II model discussion.

RTM-II allows for the oxidation of SO_2 and SO_4 by a reaction rate equation with a coefficient that varies diurnally and with latitude. Further, the model simulates a complex set of interrelated physical processes. It allows for absorption by hydrometeors (SO_2), condensation (SO_4), wet deposition, liquid phase oxidation (SO_2 to SO_4 in hydrometeors), and dry deposition. Dry deposition depends on surface roughness, a Reynolds number appropriate to the flow in the roughness layer and the ratio of the kinematic viscosity of air to the molecular diffusivity of the pollutant gas. A deposition velocity formulation is used with allowance for vegetation type and diurnal variation. Wet deposition is driven by the amount of rainfall (not an input for the present study because of the use of passive tracers). Again, diurnal variations affect the choice of coefficients in the rate equations.

The input meteorological data required for the model are exactly those required of the MESOPAC (for SRP) and MESOPAC II (for Oklahoma) meteorological preprocessors. Meteorological input data to the RTM-II model includes hourly gridded values of average wind speed and direction (averaged over 1500 meters for the Savannah River Plant cases and over the mixing height for Oklahoma), mixing height, horizontal diffusion coefficient (adjusted within user-supplied upper and lower limits), region top (mixing height plus 100 meters) and exposure class (from which stability is inferred). Both the MESOPAC and MESOPAC II models compute gridded values of stability class; the exposure class is actually only a means by which stability classes are re-created in the RTM-II model. Hourly emissions data are also required over the study period in the model (as modified by ANL/UI, see Appendix C) whereas the original model only allowed for a set pattern of diurnal variation with only the total daily emissions allowed to vary day by day.

SECTION 3

THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA BASES

3.1. INTRODUCTION

This section provides a summary of the data available as part of the Oklahoma and Savannah River Plant data bases. The field experiments are first described along with the data that are relevant to this model evaluation program. Next, the development of the modelers' data base for the Oklahoma and Savannah River Plant experiments is presented. The modelers' data base represents the ensemble of source, receptor, and ambient meteorological (surface and upper air) data that were required by the eight models. By receptor data are meant receptor locations; actual measured ground-level concentrations were listed in the data reports and had been entered in a special file for use by the statistical and graphical evaluation programs.

The largest portion of each data base is the meteorological data acquired during the experiments. It is helpful to understand the meteorological data requirements of each model before a presentation is made of the data bases themselves.

3.2. METEOROLOGICAL DATA REQUIREMENTS OF THE MODELS

All of the models under evaluation compute hourly varying wind fields and mixing heights. To provide input for the calculation of such quantities, the models require some or all of the following: surface data (typically at NWS stations), twice daily rawinsonde profiles, NCC mixing heights (MTDDIS only), and surface land use data. The ARRPA model was an exception to this pattern in that it used the wind field as computed by the NWS Boundary Layer Model. A typical user acquires the wind field already computed by NWS on a magnetic tape from either the National Weather Services Techniques Development Laboratory or from the Air Quality Branch, Tennessee Valley Authority. Table 3-1 presents a summary of the meteorological needs of each of the models.

Table 3-1. Summary of Meteorological (and Land Use) Data Required by the Eight Long-Range Transport Models

Model	Meteorological Preprocessor Used	Surface Data	Twice Daily Rawinsonde Profile	Input NCC Mixing Heights[a]	Land Use Data
MESOPUFF	MESOPAC	No	Yes	No	No
MESOPLUME	MESOPAC	No	Yes	No	No
MSPUFF	MSPACK	Yes	Yes	No	No
MESOPUFF II	MESOPAC II	Yes	Yes	No	Yes
MTDDIS	MTDDIS	Yes	No	Yes	Yes
ARRPA	BLM/MDPP	No[b]	No[b]	No[b]	No[b]
RADM	MESOPAC II	Yes	Yes	No	Yes
RTM-II	MESOPAC (SRP) MESOPAC II (OKL)	Yes[c]	Yes	No	Yes[c]

[a] All models compute mixing heights internally to the codes. Only MTDDIS requires mixing heights as a starting point for its own hourly mixing height calculation.

[b] The ARRPA model is driven by BLM-generated fields of wind components (u, v, and w), potential temperature (θ), surface-layer friction velocity (u_*) and surface-layer inverted Monin-Obukhov length (L^{-1}). BLM forecasts are based on initialized fields that are dependent on surface and rawinsonde observations. BLM output is archived by TVA and the NWS.

[c] For the Savannah River Plant experiment, RTM-II uses a wind field produced by MESOPAC, which does not employ surface data; for the Oklahoma experiment, RTM-II uses the MESOPAC II wind field which does take surface winds into account.

An agreement was made between ANL, EPA and the modelers that meteorological tower data, when available, were to be used to provide surrogate surface stations. Data not available from each tower directly were supplied from the nearest NWS station. Wind speed and direction were always obtained from the tower either from a data recording level at 10 m or by using a "1/7"-power law to extrapolate down to 10 m from the lowest available tower wind values. Only if temperature were actually measured at the 10 m level on the tower was the surface temperature for the surrogate surface station obtained from the tower. Otherwise, the value at the nearest surface station was used. No downward extrapolation of temperatures to the 10 m level was undertaken from elevated locations on the tower. Employing this method, more "surface stations" were available to provide surface wind values to the wind field models. As a result, more refined wind field predictions should have resulted. It is expected that a user (typically a source or EPA) would carry out this same procedure in a regulatory setting.

3.3. OKLAHOMA FIELD EXPERIMENTS[4]

3.3.1. Description of Experiments

A long-range tracer experiment was conducted on July 8, 1980 with the release of the perfluorocarbon, PMCH, from the National Oceanic and Atmospheric Administration (NOAA) National Severe Storms Laboratory (NSSL) at Norman, Oklahoma. Samplers were deployed to measure tracer concentrations along arcs at 100 km and 600 km north of the release point. A second experiment was conducted on July 11, 1980 with samplers located only along the 100 km arc. Each experiment involved the release of the perfluorocarbon tracer over a 3-hr period with concentrations measured 100 km downwind. In the primary experiment, the perfluorocarbon was measured at a distance of 600 km as well as 100 km. The region of plume dispersion to the north and northeast of the release site is typical of the Great Plains, including minor river valleys.

For the July 8 experiment, the PMCH tracer was released over a 3-hr period from 1900 to 2200 GMT (1400-1700 CDT) from an open field. The release nozzle was about one meter above ground level. The flow rate was monitored to assure a nearly constant release rate. The amount of perfluorocarbon released was calculated to produce concentrations well above the detection limit at the 600 km sampling arc. The background level of PMCH is extremely low (2.4 parts per 10^{15}) and also, very few points of poor data were identified from the measurements.

For the 100 km sampling arc, thirty sampling sites were selected at 4-5 km intervals (see Figure 3-1). Only seventeen samplers were available, so, based on expected winds after release, the sequential samplers, known as Brookhaven Atmospheric Tracer Samplers (BATS), were placed at sites 12-28 only. The tracer release began at 1900 GMT (1400 CDT) and the samplers were set to take ten 45-minute samples starting at 2100 GMT, before the tracer was expected to arrive.

The 600 km sampling arc consisted of 38 sampling sites through Nebraska and Missouri (see Figure 3-2). These sites represent locations of the NOAA National Weather Service substation network. All samplers on that 600 km arc started automatically at 0800 GMT (0300 CDT) on July 9. Of these 38 stations, valid data were collected for all but three sites (numbers 14, 21, and 22).

On the evening of July 8, based on the latest wind data and forecasts, the samplers were deployed to the sites indicated by double circles in Figure 3-2. Five sequential samples were taken at these locations at 3-hr intervals beginning at 1100 GMT (0600 CDT) on July 9.

3.3.2. Ground Level Tracer Observations

The sampling results for the July 8 experiment appear in Figures 3-3 to 3-5. The sampling sites are plotted as a function of azimuth from the release site. The perfluorocarbon (PMCH) concentrations for the first 6 sampling periods are shown in Figure 3-3 for the July 8 experiment. The peak plume concentrations along the 600 km arc arrived at about the same time sampling commenced at that arc. The entire record of PMCH concentrations at all sites on the 600 km arc is shown in Figure 3-4. The initial plume probably arrived at all sites on the 600 km arc just before sampling began at 0800 GMT on July 9. Plume passage had a duration of about 15 hours before background levels were seen at all locations. Background concentrations are seen for the next 15 hours, whereupon the July 11 (1400- 1700 GMT) samples show a secondary

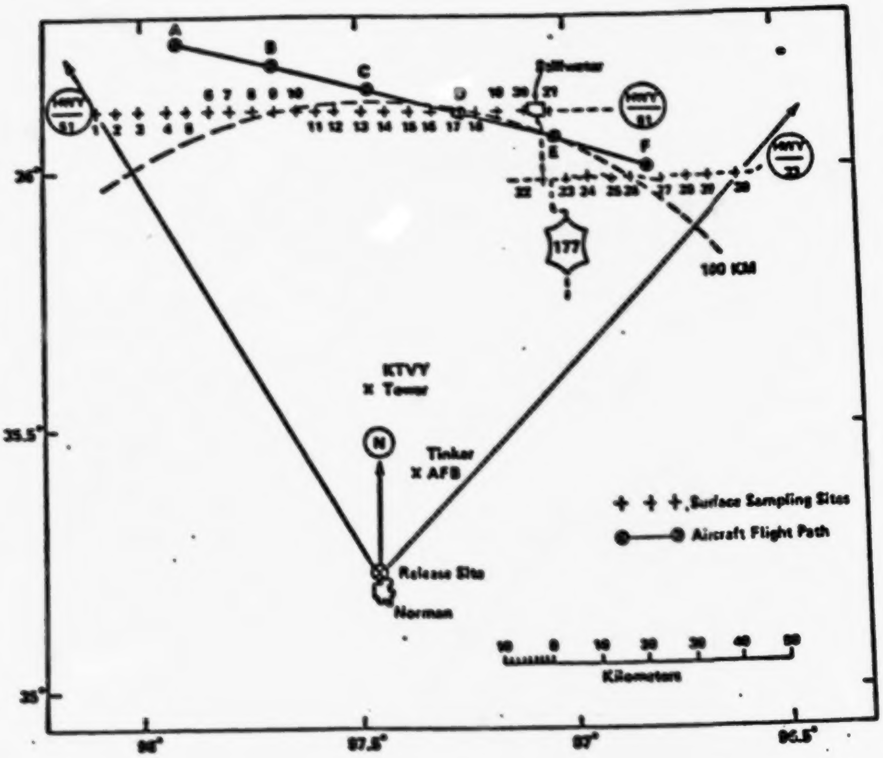


Figure 3-1. Location of the sequential air samplers (BATS) and aircraft sampling path at 100 km from the Oklahoma tracer release site.

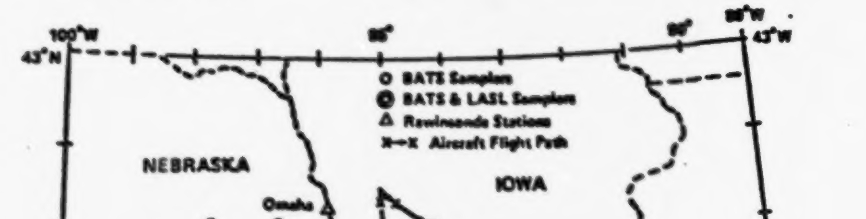


Figure 3-2. Location of sequential samplers (BATS), LASL samplers, and aircraft sampling flight path at 600 km from the Oklahoma tracer release site. The locations of rawinsonde stations are also shown.

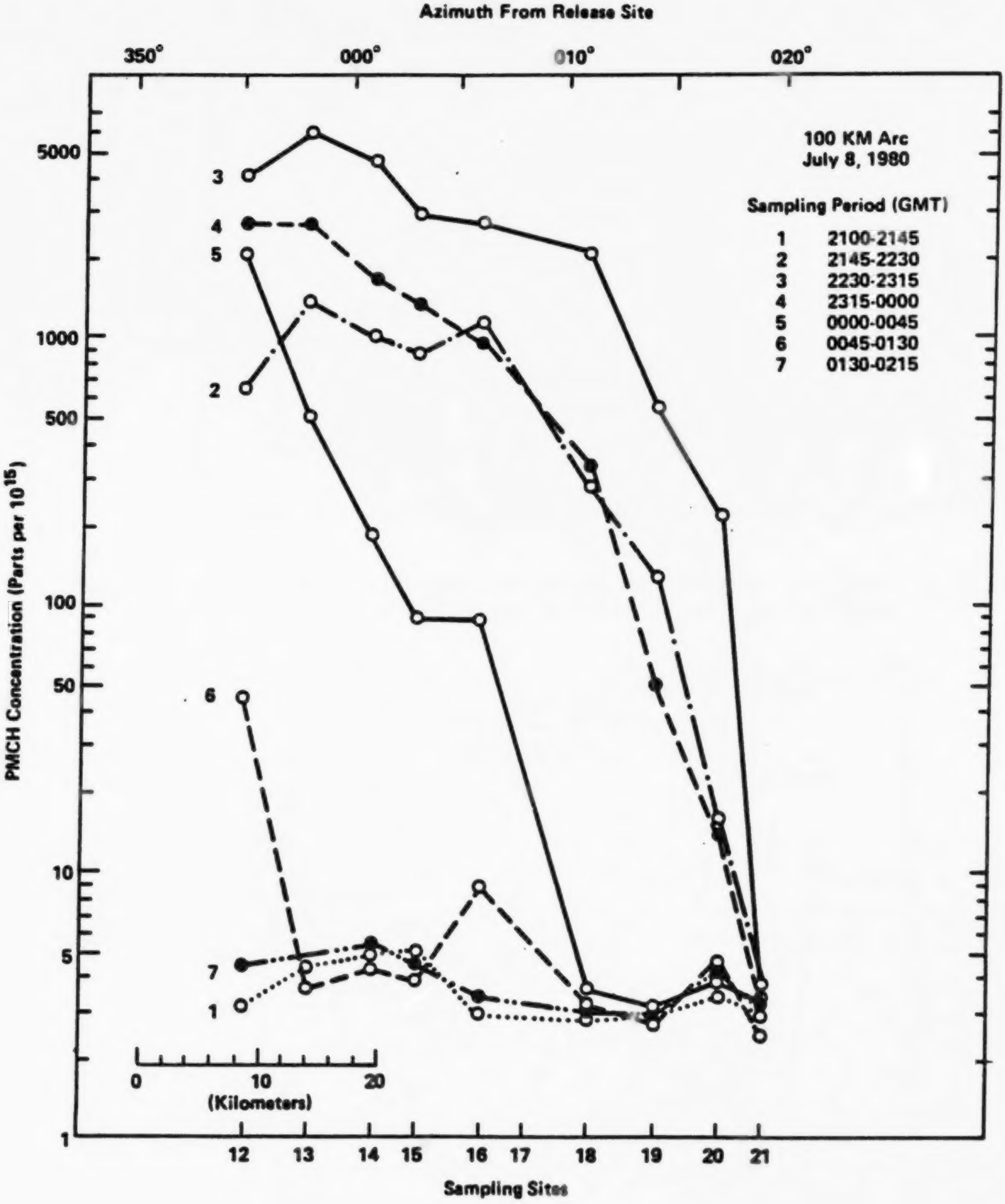


Figure 3-3. Average 45-min perfluorocarbon (PMCH) concentrations along the 100 km arc from the Oklahoma experiment No. 1 (July 8, 1980).

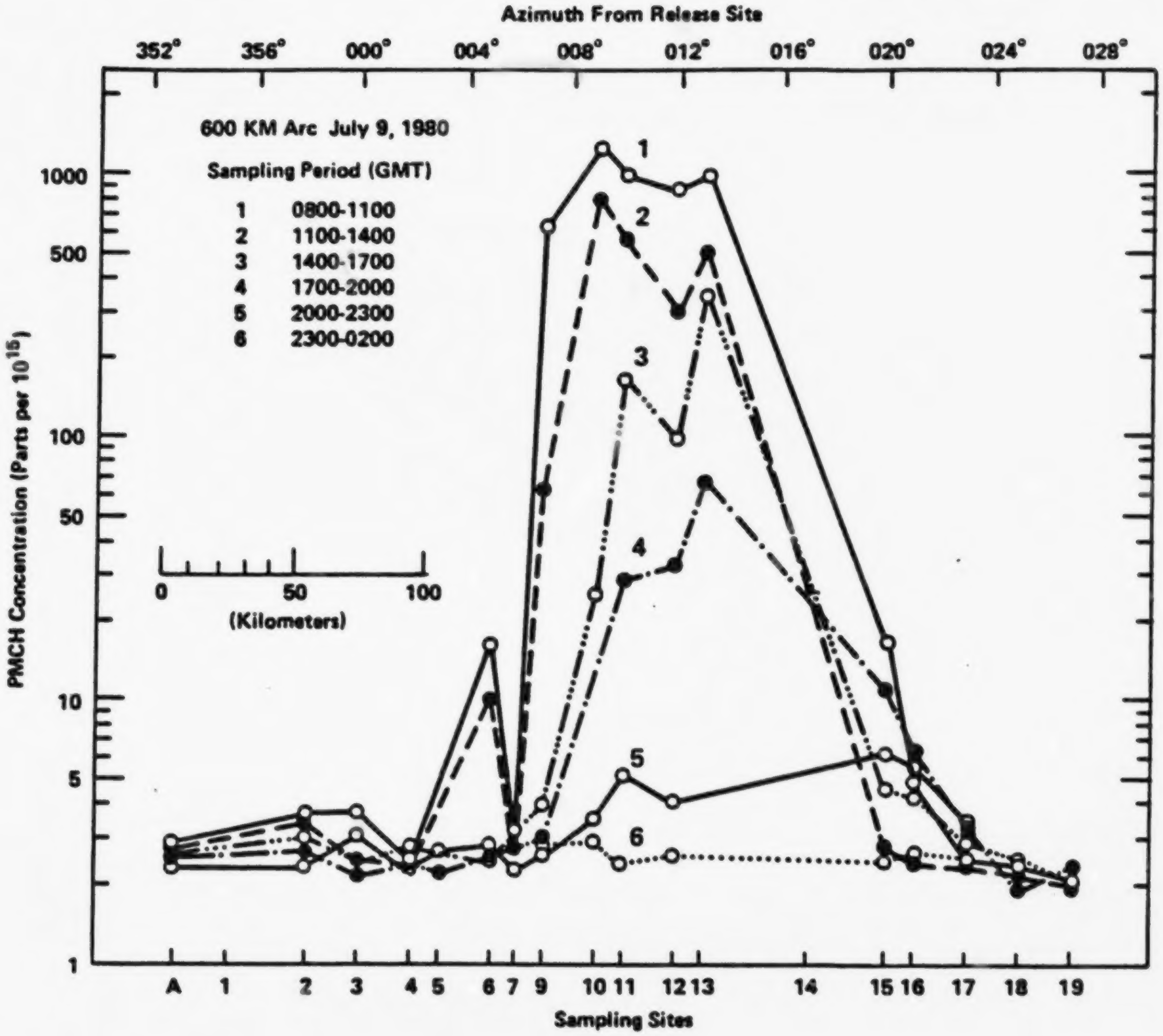


Figure 3-4. Average 3-hour perfluorocarbon (PMCH) concentration along the 600 km arc for the Oklahoma experiment No. 1 (July 8, 1980).

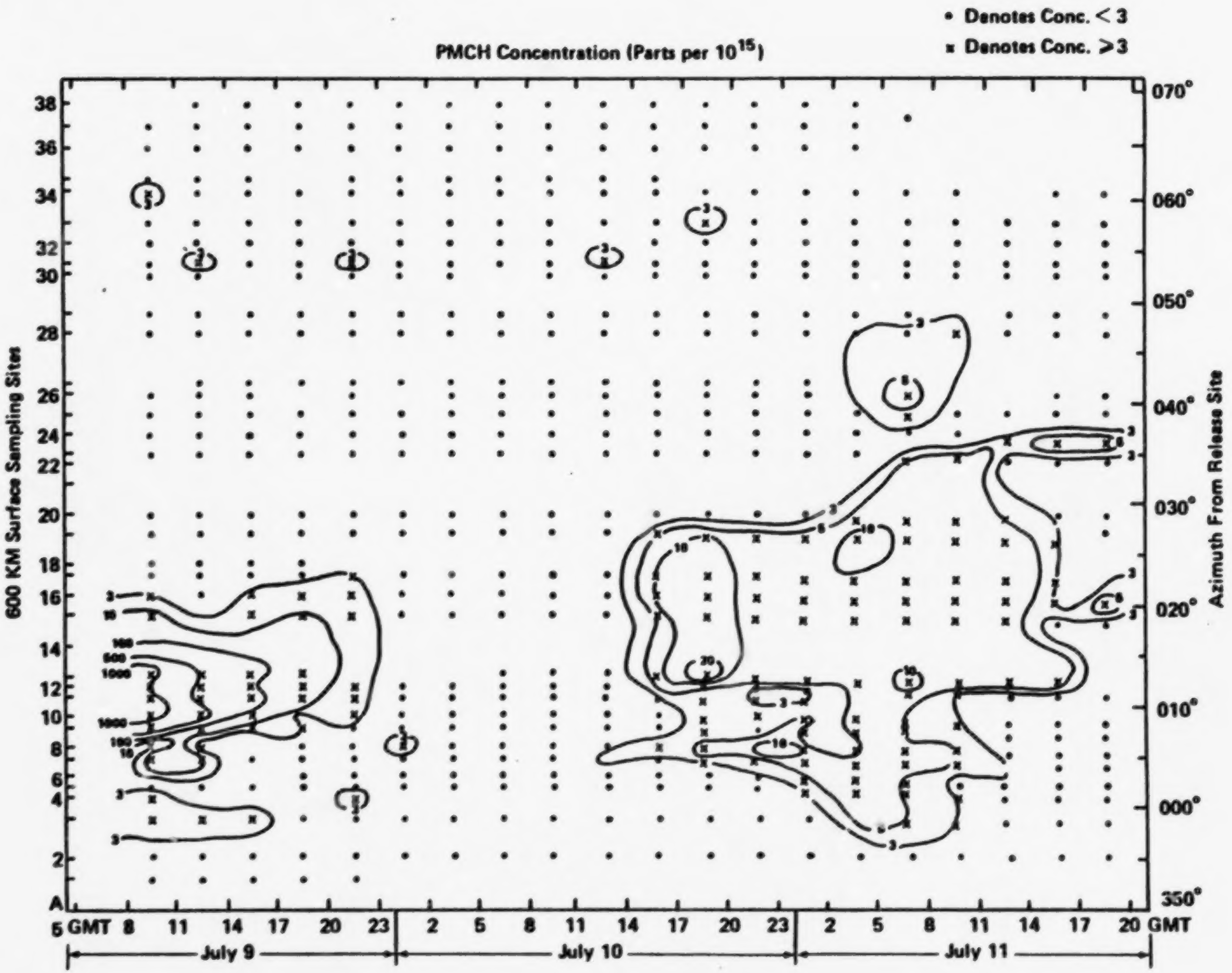


Figure 3-5. Average 3-hour PMCH concentrations along the 600 km arc for the period July 9, 0800 GMT to July 11, 2000 GMT ... Oklahoma experiment No. 1 (July 8, 1980).

plume arriving at the 600 km arc (see Figure 3-5). The maximum concentration of this secondary plume is about two orders of magnitude lower than the initial plume but they cover a much larger area. The duration of this secondary plume on the arc was about 30 hours. It is suspected that a night-time low-level jet transported a portion of the plume initially, and then the remainder of the tracer plume appeared on the arc the day following the arrival of the first portion. (It is also possible that this secondary plume is a return of the initial plume.) A discussion of the meteorological data during the first Oklahoma experiment is presented in Appendix E along with the evidence of the presence of a nocturnal jet. A second theory proposed by Dr. Ray Dickson and Dr. Gene Start of NOAA/Idaho Falls is that there was nighttime capture of the tracers on vegetation with a subsequent daytime release. This theory is not accepted by collaborator Gil Ferber of NOAA/ARL since subsequent plant studies in the laboratory have indicated no such absorption/release mechanisms. All agree that the cause of the secondary plume is uncertain and that none of the above theories has been ruled out.

A second, more limited tracer experiment was conducted on July 11, 1980. The perfluorocarbon, PMCH, was released over a 3-hr period (1900-2200 GMT) using the same release system and the same site as in the first experiment. The release amount was calculated to produce concentrations well above the detection limits at the 100 km arc.

In this experiment, sampling was done only at 100 km downwind of the release site, using the same array as in the first experiment. As in the first experiment, the BATS sequential samplers were deployed at sites 13-30 (as compared with sites 12-28 for Experiment No. 1). The tracer release began at 1900 GMT (2PM CDT) and the samplers were set to start at 2200 GMT (5 PM CDT) and take nine 45-minute samples.

The PMCH results are plotted in Figure 3-6. The initial sampling period (2200-2245 GMT) showed concentrations near background at all sampling locations. The next sampling period (2245-2330 GMT) shows concentrations at sites 14 through 24 at about 50 times background levels. During the third sampling period (2330-0015 GMT) peak plume concentrations are reached at sites 14 through 21 with decreasing concentrations to the east. During subsequent sampling periods, an orderly decrease in the PMCH concentration occurs at all sampling sites and by the eighth sampling period (0315-0400 GMT) the concentrations were approaching background levels again.

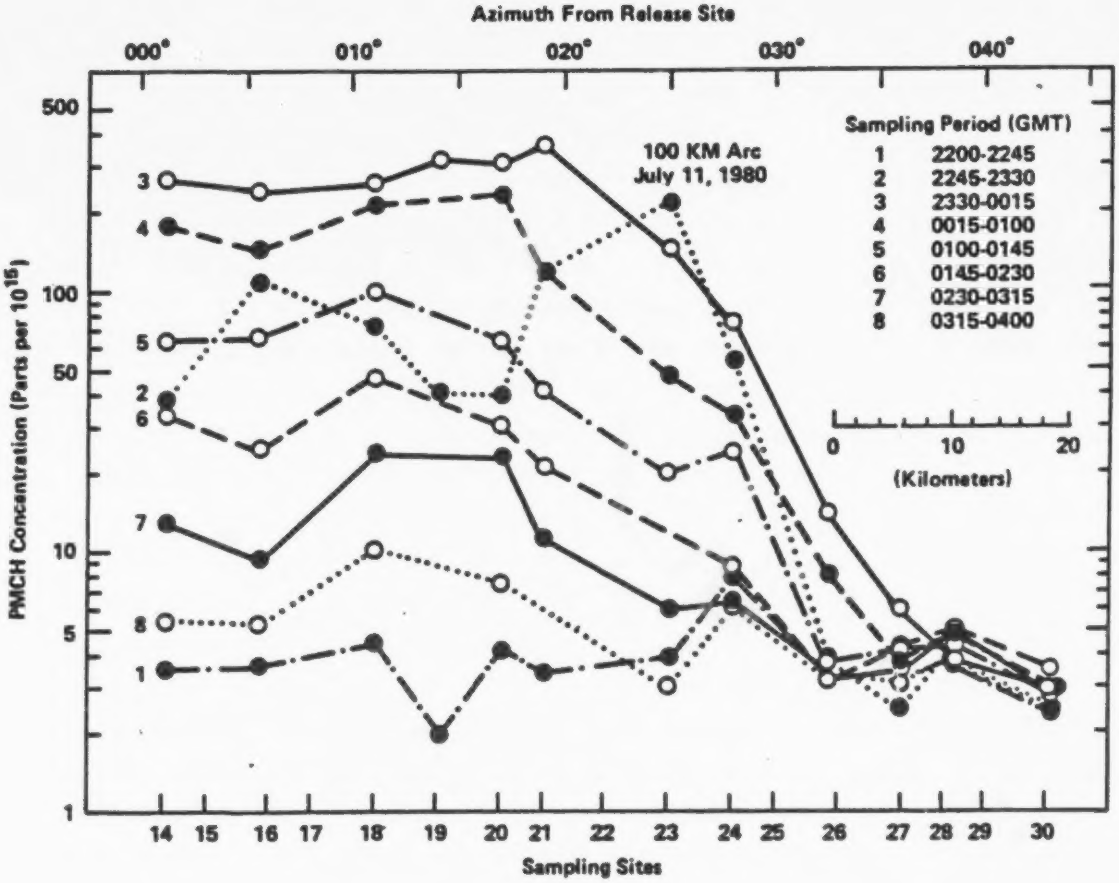


Figure 3-6. Average 45-min PMCH concentrations along the 100 km arc from the Oklahoma experiment No. 2 (July 11,1980).

Since there was no sampling west of site 14, the plume width could not be determined. Analysis of trajectories suggests that the plume did not extend much beyond site 14.[4]

These experiments provide useful case studies for verification of the eight short-term long-range dispersion models. The time-dependent tracer concentrations at a large number of points at these distances (100 and 600 km) for the July 8 experiment enhances the value of verification with these data. Unfortunately, only two experiments were made and, as a result, no long-term data base is available which would have provided an excellent opportunity for a solid statistical evaluation of the models.

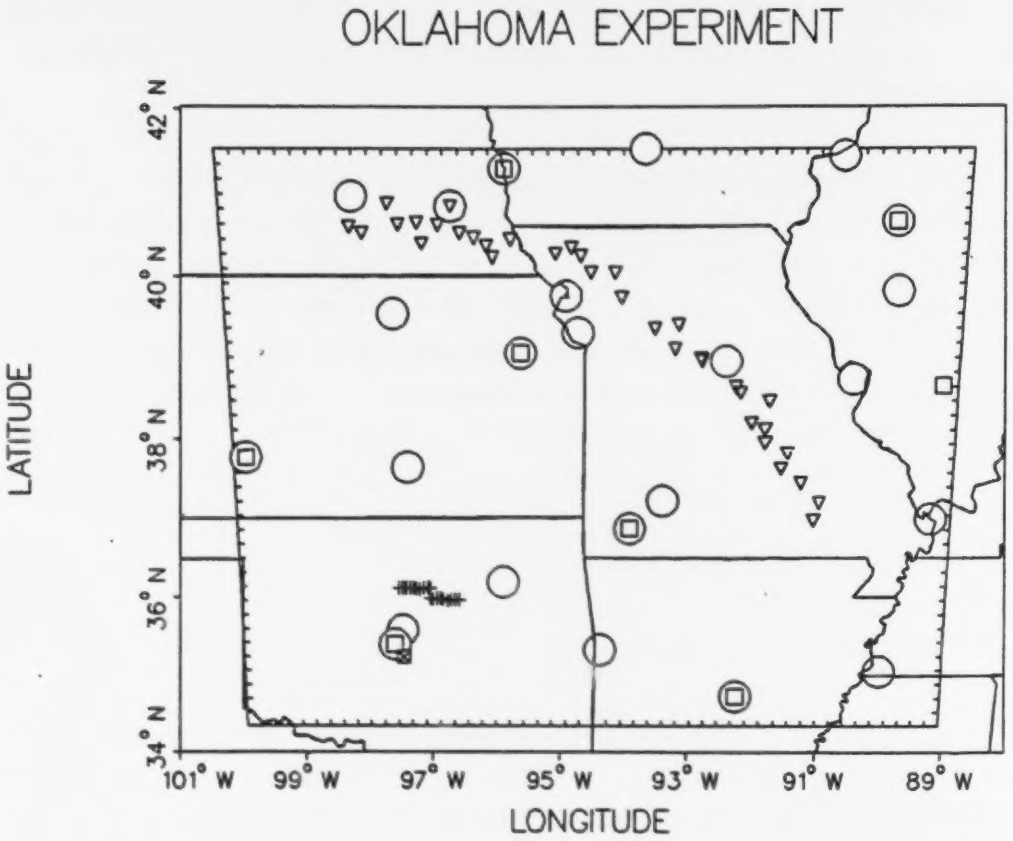
3.3.3. The Oklahoma Modelers' Data Base

3.3.3.1. The Meteorological Grid System

The grid system for the Oklahoma modelers' data base was expanded from that given in the original report by Ferber et al.[4] Figure 3-7 provides a sketch of the 24 surface stations and eight upper air stations with respect to the source and the 54 active receptors. In the original data report, 4 upper air stations were planned for use based on data listed there. In the region outlined in Figures 3-1 and 3-2, there were 14 surface stations, although none were identified in the report. However, by enlarging the region over which the meteorological wind field models would be run (to that of Figure 3-7), an additional 9 surface stations and 4 more upper air stations were included. This expansion of the wind field modeling region should lead to a more accurate prediction of the prevailing wind field and thereby a more accurate prediction of plume transport.

The meteorological grid selected for wind field predictions was 1000 km (E-W) by 800 km (N-S). Development of this grid was based on the following factors:

(a) As many surface and upper air stations as possible should be included in regions that might affect transport from the source to the receptors. In the two Oklahoma cases, prevailing southwesterly winds occurred. These factors



Source: ⊠

100 Km Arc Receptors: +

600 Km Arc Receptors: ▽

Rawinsonde Stations: □

WBAN Surface Stations: ○

Figure 3-7. Location of significant source, meteorological, and receptor sites for the Oklahoma Cases No. 1 and 2.

dictated expanding the area from the region presented in the Ferber report (Figures 3-1 and 3-2) to the area covered in Figure 3-7.

(b) Several of the models require a meteorological "cushion" of at least 3 grid points around the region of concentration calculations.

(c) It is not desirable to keep expanding the region beyond the boundaries of the active region, because the influence of a surface station or upper air station is weighted usually by the inverse square of the distance to the computational point.

(d) Long range transport models that use a grid or mesh are seldom run with more than 1200-1600 grid points. The grids used in model development and application usually range from 30 by 30 to 40 by 40 grid cells. Further, there is little point in achieving a spatial resolution that is much finer than the average spacing between surface weather stations or sources, since conditions and variations in meteorology cannot be resolved on the finer scale.

For the Oklahoma experiment, these four criteria were met easily by expanding the grid to 51 by 41 with 20 km square grid spacing. This choice yields over 2000 mesh points, and has small enough grid spacing to accommodate variations between surface stations.

As a result of this expanded meteorological base, more accurate wind field predictions result from the models, thereby permitting a fairer test of model dispersion capabilities. A total of 25 surface meteorological stations were used for the two Oklahoma tests. Of these, only Salem, Illinois had highly incomplete surface data for the study period (July 7-12, 1980) and Salem, Illinois had no surface data (only upper air data). This left a total of 24 surface stations. None of the surface station data were included in the Ferber data report.

3.3.3.2. Surface Weather Observations / Modelers' Data Base

NWS surface station data for 24 stations within the selected modeling grid region were ordered from the National Climatic Center (NCC) in Asheville, North Carolina. Five of the eight models (MSPUFF, MESOPUFF II, MTDDIS, RTM-II, and RADM) require surface data as input.

All of the surface meteorological data were acquired from the NCC in the form of photocopies of the original sheets on which full hourly data were recorded. Data from such sheets are normally entered into the NCC computer for the development of the TDF-14 and CD-144 magnetic tapes. However, these archived NCC tapes only contained 3-hourly data during the study period. By obtaining only the sheets needed, it was possible to acquire all of the hourly data. For relatively short study periods, this procedure is reasonable for any model user to follow, but for periods longer than a week, the effort and cost entailed may not be justified. The user might then have to rely on the 3-hourly data normally provided by the NCC in the CD-144 or TDF-14 format. (More recently, the NCC has been archiving hourly data.)

3.3.3.3. Rawinsonde Observations / Modelers' Data Base

Rawinsonde data from Oklahoma City, Oklahoma; Monett, Missouri; Topeka, Kansas; and Omaha, Nebraska; were presented as part of the original data base.[4] These observations were from July 8, 0000 GMT to July 12, 1200 GMT. The data included pressure, altitude (m above sea level), air temperature, dew point depression, wind direction, and wind speed. In addition, special rawinsonde observations were taken at Tinker Air Force Base about 20 km NNE of the release site, starting on the morning of July 8. These data (height, temperature, wind speed and direction) were available in tabular form.[4]

The upper air data contained in the Ferber report[4] were also augmented by obtaining from NCC the computer printouts of the original soundings for additional times (at stations reported in Ref. 4) and at all times (for stations not reported in Ref. 4) within the study period. The final study period for upper air data was selected to run from 1200 GMT on July 7 to 1200 GMT on July 12. Table 3-2 provides the timing of releases, sampling and periods of computation for Cases No. 1 and 2 for Oklahoma.

Table 3-2. Timing Schedule of Releases, Ground-Level Samples, and Computational Periods for Oklahoma Cases No. 1 and 2.

```
          TIMING OF RELEASES, SAMPLES AND COMPUTATION CASES 1 AND 2 FOR OKLAHOMA
                            (All times in GMT)
    JULY 7  |        JULY 8         |        JULY 9         |       JULY 10         |       JULY 11         |  JULY 12  |
1  1  1  2  0  0  0  0  1  1  1  2  0  0  0  0  1  1  1  2  0  0  0  0  1  1  1  2  0  0  0  0  1  1  1  2  0  0  0  0  1
2  5  8  1  0  3  6  9  2  5  8  1  0  3  6  9  2  5  8  1  0  3  6  9  2  5  8  1  0  3  6  9  2  5  8  1  0  3  6  9  2
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0

                              RRR                                                                     RRR

                                 111111111                                                                111111111

                                           6666666666666666666666666666666666666666666666666666666666666

C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C1C

                                                                       C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C2C
```

Legend: R = release period
1 = measurements at 100 km arc
6 = measurements at 600 km arc
C1 = computational period for July 8 release
C2 = computational period for July 11 release

Note: The total number of averaging periods for the Oklahoma data base is 21. The breakdown is as follows:

July 8 release (100 km arc) : 7
July 8 release (600 km arc) : 6
July 11 release (100 km arc): 8

Of the eight upper air stations employed in this model evaluation study, data from at least six are available during each 12-hour period. Table 3-3 summarizes availability and periods of record for the eight rawinsonde locations. During the Oklahoma experiment, special upper air soundings were taken at three of the eight stations to enrich the detail of the upper air data during the time period when perfluorocarbon tracer was expected to be spreading toward the samplers. It was found, however, that the extra rawinsonde soundings at the special 6-hourly intervals could not be used by the models. The models, in their present form, accept only 12-hourly rawinsonde data at all stations. As a result, the extra 9 soundings could not be used by the models.

3.3.3.4. Meteorological Tower Observations / Modelers' Data Base

Wind speed and direction are available from the KTVY tower located about 40 km north of the release site. The tower is instrumented at seven levels between the surface and 444 meters. The wind data provided at these levels were averaged over 15-minute periods. Data from the KTVY tower are available at heights of surface, 24, 45, 89, 177, 266, and 444 m. The period of measurement is from 1815 GMT July 8, 1980 to 0200 GMT July 9, 1980. No tower data are available for later periods of the July 8 experiment or for the July 11 experiment. A surrogate surface station was made from the KTVY tower data.

3.3.4. The Treatment of Missing Data

The treatment of missing data was resolved on a model-by-model basis for the Oklahoma (and SRP) data based on instructions in the user's manuals. For MESOPAC, the code handles missing data internally. MESOPAC II, however, requires all missing data (surface and upper air) to be filled in with replacement data either by (a) interpolating in time using data at the same station, or (b) by using data from a nearby station if the period of data loss was large (more than just a few hours). Subjective judgment was required for many cases of missing data for MESOPAC II in order to determine the best alternative. Judgments were made primarily on the basis of how conditions

Table 3-3. Availability and Periods of Record for the Eight Rawinsonde Locations for Oklahoma Cases No. 1 and 2.

(1) SAL is Salem, IL -- 03879
(2) UMN is Monett, MO -- 03946
(3) LTR is Little Rock, AK -- 13963
(4) OKC is Oklahoma City, OK -- 13967
(5) DGC is Dodge City, KA -- 13985
(6) TOP is Topeka, KA -- 13996
(7) PEO is Peoria, IL -- 14842
(8) OMA is Omaha, NB -- 94918

DATE GMT	DATE CST	YYJJJ GMT	TIME GMT	TIME CST	SAL	UMN	LTR	OKC	DGC	TOP	PEO	OMA	NO. TAKEN	SEQ. NO.
7/07	7/07	80189	12	06	A	A	A	A	A	A	A	A	8	0
7/08	7/07	80190	00	18	A	A	A	A	A	A	A	A	8	1
7/08	7/08	80190	12	06	A	A	A	A	A	A	A	A	8	2
7/08	7/08	80190	18	12	[ ]	A	[ ]	[ ]	[ ]	A	[ ]	A	3	3
7/09	7/08	80191	00	18	[X]	A	A	[X]	A	A	A	A	7	4
7/09	7/09	80191	06	00	[ ]	A	[ ]	[ ]	[ ]	A	[ ]	A	3	5
7/09	7/09	80191	12	06	[X]	A	A	A	A	A	A	A	7	6
7/09	7/09	80191	18	12	[ ]	A	[ ]	[ ]	[ ]	A	[ ]	A	3	7
7/10	7/09	80192	00	18	[X]	A	A	A	A	A	A	A	7	8
7/10	7/10	80192	12	06	[X]	A	A	A	A	A	A	A	7	9
7/11	7/10	80193	00	18	[X]	A	A	A	A	A	A	A	7	10
7/11	7/11	80193	12	06	[X]	A	A	A	A	A	A	A	7	11
7/12	7/11	80194	00	18	[X]	A	A	A	A	[X]	A	A	6	12
7/12	7/12	80194	12	06	[X]	A	A	A	A	A	A	A	7	13
			TOTALS		3	14	11	11	11	13	11	14	88	

A = Available and included in the data base
[ ] = Not available
[X] = Not available, but filled in from nearest neighbor rws for some models

There are 88 available soundings out of 112 possible. With the "fill-ins" there are still 88 soundings; the soundings at 06Z and 18Z cannot be used by any of the models.

were changing for the variable of interest (e.g. wind speed, wind direction, opaque cloud cover, etc.) at the station and for nearby stations. For the MSPACK preprocessor, missing data are treated internally within the code.

The following additional adjustments were made due to missing data based on the requirements of the meteorological processors:

(a) At Salem, Illinois, eight of the eleven total rawinsonde soundings were missing. MESOPAC II requires a continuous record from any rawinsonde site that is utilized. The three models that employ MESOPAC II are MESOPUFF II, RADM, and RTM-II (Oklahoma data only). For MESOPAC II therefore, it was necessary to augment the three available Salem soundings with eight soundings from Peoria, Illinois, the nearest station that had upper air data.

(b) At Topeka, Kansas, one upper air sounding was missing, and for MESOPAC II based models, it was replaced by the nearest one from Omaha, Nebraska.

(c) The addition of one surface station was made. This station was created from the 10-m level winds of the KTVY tower and the temperature and cloud data from the nearby Oklahoma City surface station. Only nine hours of data were obtained from this station, since winds were measured for only a 9-hour period between July 8 and 9, 1980.

(d) In the surface data sheets obtained from NCC, the lowest layer cloud height was only recorded every three hours, although the ceiling was noted every hour at most stations. To provide hourly values of lowest layer height, a modified rule was followed. A given lowest layer height was used for the next two hours unless additional cloud height data recorded in the ceiling column indicated an intermediate change. (The ceiling column showed several layers, and usually reflected the presence of lowest layer clouds in agreement with the lowest layer height column at three-hourly intervals.)

3.3.5. The Source Release / Modelers' Data Base

One further issue was the nature of the tracer release in the Oklahoma experiment. All of the models expect the emissions from a point source stack with buoyancy or from an area source. At Oklahoma, the neutrally buoyant tracer was released into the prevailing wind from a nozzle at a height of 1 m above the ground. The area over which the release was made is far too small to fit the area source category, and normal type exit conditions did not fit the tracer source configuration. Exit conditions were devised that gave a large densimetric Froude number and a small plume rise—on the order of 10 m under most conditions. This allows the models to calculate a plume rise value without numerical difficulties. As recommended by the model developers, each code was modified by resetting the plume rise to 1 m above ground after the artificial plume rise was calculated.

3.4. SAVANNAH RIVER PLANT KRYPTON-85 EXPERIMENTS[5]

3.4.1. Description of Experiments

The second data base used for model testing represents observations taken routinely in the vicinity of the Savannah River Plant during the period October 1976 through July 1977. Krypton-85 is a noble gas which is released to the atmosphere during the chemical separation of nuclear fuel from target material at the Savannah River Plant (SRP) in Aiken, South Carolina. The release comes from two 62 m stacks located 4 km apart. The release is strongly time dependent. The krypton-85 release rate is based on calculations which are estimated to be accurate to within about a factor of two for an individual hour and to within 10% for daily averages.

The terrain within 150 km of the SRP is represented by gently rolling hills ranging in elevation from 150 m above sea level to the northwest to about 25 m toward the southeast. The SRP is covered with mixed hardwood and pine forests; the surrounding area consists of equal amounts of mixed forests and cleared farm land.

3.4.2. Ground-Level Krypton-85 Observations / Study Periods

The krypton-85 samplers were located at 13 sites surrounding the Savannah River Plant. The nearest sampler was 28 km from the stacks and the farthest was 144 km. Twice-daily samples representing 10-hr averages were used in this study for the testing of models.

Fifteen data sets were chosen for model evaluation. The fifteen were the "standard" data sets defined by the Savannah River Laboratory (SRL) for use in the Model Validation Workshop held in Hilton Head, South Carolina in November 1980. Table 3-4 lists the time periods over which 10-hour samples were collected for inclusion in the 15 data sets. At least three data sets are available for each season of the year. Each data set represented roughly 3-6 days of between one to twelve 10-hour sampling periods. In general, 1-1.5 days of plume transport calculations were carried out in advance of the start of ground sampling measurements to assure that the plume is sufficiently transported downwind to be predicted accurately for the same periods as the samples were measured.

The fifteen periods were defined by SRL after careful evaluation of all the data measured during the 1975-1977 period of study. The quality of meteorological and ground-level data, and the presence of sufficient above-background data readings were key considerations in the choice of these data sets by SRL. Table 3-4 lists the sampling period and computational period for each of the 15 SRP data sets.

3.4.3. The Savannah River Plant Modelers' Data Base

3.4.3.1. The Meteorological Grid System

The locations of the rawinsonde stations, sampler locations, surface stations, and meteorological tower locations are presented in Figure 3-8. Their locations are superimposed on the grid system used in model testing. The meteorological grid is 33 by 56 with 10 km by 10 km grid cells (total grid size is 320 km by 550 km). This is slightly larger than the grid system used in the previous work with MESOPUFF and RTM-II (see Ref. 23, Appendix A).

Table 3-4. Sample Collection Periods and Model Calculational Periods for 10-Hr Samples ... Representing 15 Data Sets for the Savannah River Plant Krypton-85 Data Base.

	Sample Collection Periods				Model Calculation Periods			
	Start		End		Start		End	
Case No.	Hr	Date	Hr	Date	Hr	Date	Hr	Date
1.	2200	10-05-76	1200	10-06-76	0000	10-05-76	0000	10-07-76
2.	2200	10-14-76	1200	10-16-76	0000	10-14-76	0000	10-17-76
3.	0900	10-29-76	0700	10-30-76	1200	10-28-76	1200	10-30-76
4.	1000	11-18-76	0800	11-20-76	1200	11-17-76	1200	10-20-76
5.	1000	02-02-77	0800	02-04-77	1200	02-01-77	1200	02-04-77
6.	2200	02-16-77	0800	02-19-77	1200	02-15-77	1200	02-19-77
7.	1000	02-22-77	0800	02-23-77	1200	02-21-77	1200	02-23-77
8.	2200	04-05-77	0800	04-09-77	1200	04-04-77	1200	04-09-77
9.	1000	04-11-77	0800	04-16-77	1200	04-10-77	1200	04-16-77
10.	2200	04-17-77	2000	04-22-77	0000	04-17-77	0000	04-23-77
11.	0900	04-27-77	0700	04-29-77	1200	04-26-77	1200	04-29-77
12.	0900	07-11-77	0900	07-12-77	1200	07-10-77	1200	07-12-77
13.	0900	07-15-77	0700	07-16-77	1200	07-14-77	1200	07-16-77
14.	0900	07-18-77	0900	07-20-77	1200	07-17-77	1200	07-21-77
15.	0900	07-25-77	0700	07-27-77	1200	07-24-77	1200	07-27-77

Note: The total number of 10-hour averaging periods for the SRP data base is 65. The breakdown is as follows into subcases:

Experiment	No. of Subcases
1	1
2	3
3	2
4	4
5	4
6	5
7	2
8	7
9	10
10	10
11	4
12	2
13	2
14	5
15	4

This region and grid interval choice satisfies the four criteria listed in the discussion of grid choice for the Oklahoma data previously.

3.4.3.2. Surface Weather Observations / Modelers' Data Base

Meteorological data for the krypton-85 Savannah River Plant experiments had been placed on magnetic tapes for the region within and surrounding the Savannah River Plant by the Savannah River Laboratory. These meteorological data bases will be referred to below.

The Savannah River Plant data base contains surface weather observations at stations between 100 W and 60 W, 50 N and 20 N. These weather observations were originally provided to SRL on magnetic tape from the NOAA National Climatic Center. About 600 surface weather stations report each hour in this general area. As many as 35 items, mostly meteorological variables, are recorded for each station. Of these variables, wind speed, wind direction, air temperature, ceiling, total opaque sky cover, and height of lowest cloud layer were used in the data base. Figure 3-8 identifies the surface stations within 200 km of the SRP source. Surface weather stations located between 86 W and 77 W, 37 N and 30 N are included in the four tapes which contain the SRP data base. Ten parameters per station had been added to the original four SRP tapes by the Savannah River Laboratory (SRL). These ten parameters are the World Meteorological Organization (WMO) block station number, station longitude and latitude, station elevation above mean sea level, wind direction, wind speed, station pressure, dry bulb temperature, dew point depression, and the previous 6-hour precipitation amount. In addition, the Pasquill stability category (A to G represented by 1 to 7) as defined by Turner has been added to each surface observation.

One problem area that required work was the need to append the cloud data (ceiling, lowest layer height and total opaque sky cover) to each hour for each surface station. Such cloud data were not on the SRL meteorological tapes. To acquire the additional cloud data, the full surface meteorology tapes containing data from 600 stations, including the 19 in the SRP study region, were obtained from Dr. Roland Draxler of NOAA/ARL. The meteorological records for each station at each hour as contained in the original SRP data base were augmented with the needed cloud data to produce new surface

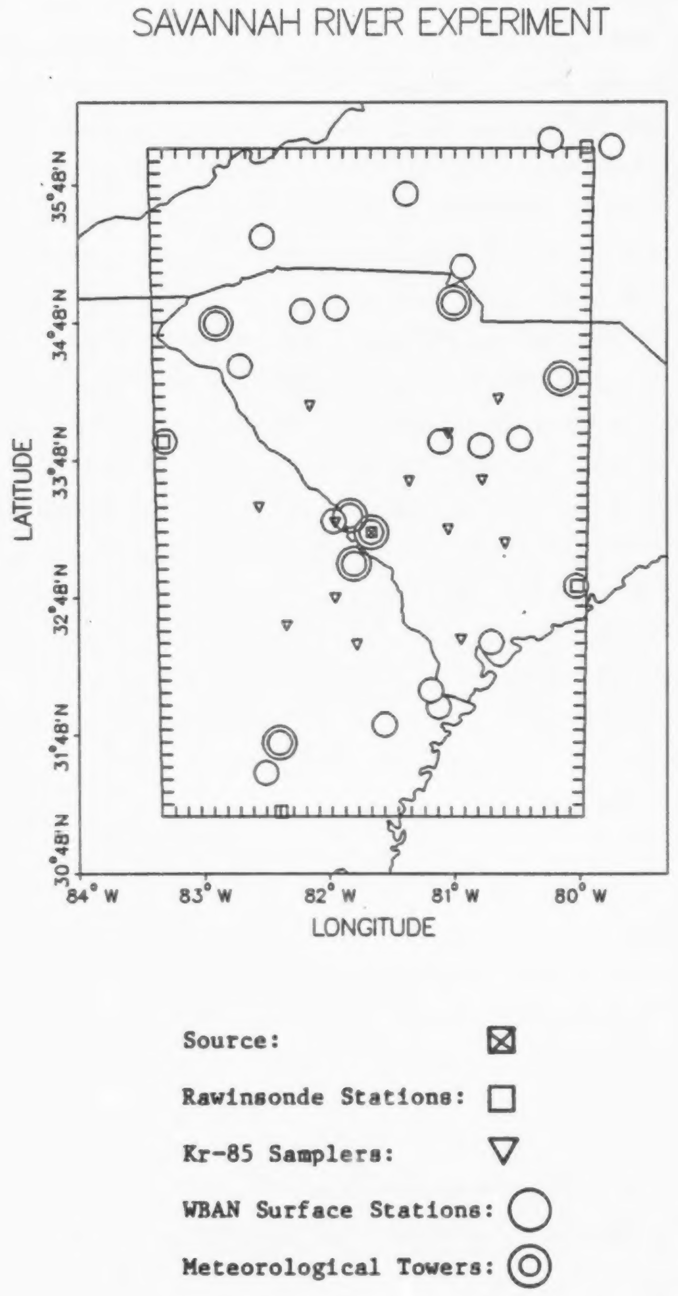


Figure 3-8. Location of significant source, meteorological, and receptor sites for the Savannah River Plant data base (includes final region choice; inner box is model solution region).

meteorological records. A number of models use cloud cover and ceiling, but lowest layer cloud height is used only by MTDDIS. The Air Force surface station data on these tapes define lowest layer cloud height differently than they are defined by the NWS. A consistent interpretation of these NWS and Air Force definitions was developed which was later approved by Dr. I-Tung Wang of Rockwell International (for MTDDIS). For the NWS data, which was the prototype data for the development of MTDDIS, the lowest layer cloud height is usually reported every 3 hours, while ceiling and total opaque sky cover are reported hourly. The lowest layer height for non-reported hours was obtained by using the last reported measurement for two more hours. However, the ceiling entry often recorded multiple layers, including the lowest layer. Where available, these data were used in preference to using the last 3-hourly lowest layer value. In the Air Force data, "lowest layer" is defined as clouds lying from the ground to a fixed height, while in the NWS data it is defined as the lowest layer present. The Air Force data reports lowest layer, middle layer and high clouds. The smallest of the numbers recorded at an hour for the lowest layer and the largest for the ceiling were used, thus achieving the closest commonality possible between the definitions.

3.4.3.3. Rawinsonde Observations / Modelers' Data Base

Upper air soundings in the SRP data base are available from four stations at 12-hour intervals: Waycross and Athens, Georgia; Greenville and Charleston, South Carolina. Temperatures were linearly interpolated to a wind level where no corresponding temperature existed. Temperatures at levels where no winds were available were excluded.

Related to the rawinsonde data is the need for NCC mixing heights as input to the MTDDIS model. These mixing heights were obtained for the Charleston, South Carolina rawinsonde station. (For the distance scale of this SRP data base, the mixing heights for only one station were adequate for input to MTDDIS.)

3.4.3.4. Meteorological Tower Observations / Modelers' Data Base

An additional problem area related to whether the available meteorological tower data should be used or not. It was decided to use these data by adding seven surrogate surface locations to the 19 actual surface stations.

Meteorological data collected at power plant sites by seven meteorological towers in the area near the Savannah River Plant were provided by local utility companies (Carolina Power and Light Co., Duke Power Co., Georgia Power Co., South Carolina Electric and Gas Co.) as part of the SRP data base. The Newberry, SC tower never recorded data during the experimental period (only before), and the Southport, NC tower is outside the meteorological grid. In effect, there were only five. With the addition of two "special towers" described below, that totals seven for use as surrogate surface stations. The data from the seven meteorological towers were reformatted to simulate seven surface stations. Added to the 19 actual surface stations in the grid, the total becomes 26.)

In most cases the measurements on the five power plant towers included hourly wind speed and wind direction at one to three levels and ambient temperature, dew point, vertical temperature gradient, precipitation and solar radiation when available. All power plant towers contained surface level data.

In addition to the five remaining power plant towers, meteorological data from the WJBF television tower, 21 km from the SRP source, are included with the other data as the sixth tower. The TV tower is instrumented at seven levels between 2 and 335 m above ground with temperature sensors and turbulence-quality wind sensors. Data at three levels (10 m, 91 m and 243 m) were averaged over 15 minute periods and tabulated at one-hour intervals by SRL.

Adjacent to the main SRP operating areas, seven on-site towers with a wind sensor at 62 m only were located in pine forests within a 10 km radius of the SRP source. The data from these were combined by SRL to make a seventh tower due to (a) the close proximity of the on-site towers to each other, and (b) the problems with the availability of data from all stations at the same time. To this end, the 15 minute average wind speeds and directions were tabulated at one hour intervals. Because any individual 62 m tower usually did not have a continuous record, an hourly arithmetic average of the wind velocity from all on-site towers was then calculated from the available data at one-hour intervals. This arithmetic average was assumed to be representative of the wind conditions at 62 meters (source area stack height) at a location midway between the two source areas, F and H. The seven towers at the SRP are shown in Figure 3-8. In summary then, seven surrogate surface stations were created for the Savannah River Plant data sets: five from off-site power plant towers, one from the WJBF tower and one from the combined seven on-site towers at the SRP.

3.4.3.5. The Source Release / Modelers' Data Base

At the Savannah River Plant there were two identical stacks emitting krypton-85 located about 4 km from each other. The emissions data that were available from the Savannah River Laboratory provided only the sum of the emissions from the sources, and individual source emission data were not available. Further, in light of the 10 km grid spacing and the fact that the nearest sampler location was 28 km from the sources, the two sources were replaced with a single effective source whose diameter was enlarged from the actual stack diameter by a factor equal to the square root of two. It would have been desirable to use a two-source simulation with a 4-km separation. However, discussions with Dr. Allen Weber of the Savannah River Laboratory revealed that the separation of emission rates between the two sources is classified information. The emissions sum was carefully evaluated before being unclassified to assure that it would not be possible to separate out the two individual emission rates. The split of krypton-85 emissions between the two site areas (one stack for each area) of SRP is not to be revealed. The only calculations done previously with the separate source emission rates were made for samplers very close to the sources. Those calculations are classified as well but it is known, however, that the two stacks do not have the same emission rates. It was felt that given combined emissions data, the single effective source was a consistent and better choice than the use of half the total for each of the two separate sources. Any errors introduced by this combined source schematization should occur only at the nearest samplers.

The actual exit velocity was available and was used, but the exit temperature was not measured and varies from a few degrees below to a few degrees above ambient temperature. A uniform value of 1 degree Centigrade above ambient was used for exit temperature to provide a slightly buoyant plume. Although the plume rise calculated using these values will differ somewhat from the actual rise (within the validity of the formulas used for plume rise) because of ignorance of the actual positive or negative buoyancy of the plume, the differences appear to be less than 10% of the rise. Further, the nearest sampler was 28 km distant, and under most conditions the plume would occupy most of the mixing layer by the time it had transported to the first sampler. As a result, this assumption is not considered a significant source of error.

SECTION 4

OPERATIONAL EVALUATION BASED ON MODEL/DATA COMPARISONS

4.1. INTRODUCTION

This section provides a summary presentation of the model/data comparisons made with the eight models for the Oklahoma and Savannah River Plant data bases. An operational evaluation of the models is presented using several methods of comparing model predictions with the field data. In Section 5, a diagnostic evaluation will be presented which is aimed at tracing back the causes of model/data discrepancies to model assumptions. Additional methods of testing model performance are presented there to provide added insight into the causes of model/data discrepancies.

The operational evaluation relies mainly on the use of the American Meteorological Society (AMS) statistics[3] recommended to EPA by the AMS for the evaluation of the performance of air quality models. Additional quantitative insight into systematic differences between model predictions and data, is obtained by examining a set of graphs for each model, which compare predicted and observed concentration values in several special ways. This section will provide only the highlights of the numerous graphs and statistical tables that were generated during the preparation of the model/data comparisons. A more complete summary of the tables and figures used in the evaluation appears in Appendices D and E.

In Section 4.2 the results of the AMS statistics are applied to each of the eight models. Tables of the various statistical measures are presented separately for the Oklahoma and the Savannah River Plant data sets. The AMS statistics provide a method of model evaluation commonly used by the EPA in evaluating air quality models. Additional graphical methods of presenting model/data comparisons have been developed and are presented in Section 4.3. These graphs primarily confirm the evaluations based on the AMS statistics, but provide a visual means of examining the predictions of each model as compared with the data. They also provide some additional insight into the performance of the models.

One class of the AMS statistical measures in Section 4.2 involves a comparison of the set of predicted concentrations to the set of observed concentrations, regardless of any time relationship between them. These measures involve points not "paired in space and time." A second class involves "points paired in space and time," which means that predicted and observed values are compared at the same sampler for the same averaging period. In general, the set of predicted and observed values was found to agree best when they are not paired in space and time because the most common inaccuracy in the predicted plume patterns is a shift in the overall pattern away from the observed one either in ground position or in time of arrival or in both respects. In Section 4.3, all of the graphs and tables involve points paired in space and time.

4.2. APPLICATION OF AMS STATISTICS TO THE OKLAHOMA AND SRP DATA BASES

4.2.1. Introduction

The primary means of evaluating model performance in this report is through the use of the American Meteorological Society statistics, prepared from predicted and observed concentration values for each model on a paired basis. These statistics are described in general form in Reference 3. The AMS workshop recommended that performance evaluations be based on comparisons of the full set of observed/predicted data pairs, of the highest observed and predicted concentration per event (e.g., 1, 3, or 24 hour time period) and of the highest N values (unpaired in space or time). In addition, if the data are sufficient, comparisons of observed and predicted concentrations may be carried out on data subsets representing individual monitoring stations or selected meteorological conditions.

The guidance provided in the Woods Hole Workshop report[3] is quite general and broad. The statistics were aimed at air quality model evaluation exercises in general; no specific guidance was given on the application to any model category such as long-range transport models. However, much has been learned in the last few years regarding application of the generalized statistics to the evaluation of the rural[48], urban[49], and complex terrain[50] models. Much of that experience is transferable and was used in the evaluation of long-range transport models with the Oklahoma and SRP data sets. The format of the discussion below closely follows that of the previous EPA/TRC evaluations[48-50] using the AMS statistics. In this way, it would be possible to compare easily the performance of models of different categories, e.g. urban and long-range transport.

4.2.2. Statistical Data Sets for Comparison of Observed and Predicted Concentrations

Table 4-1 summarizes the data sets to which the AMS statistics in their most general form may be applied. The data sets listed in Table 4-1 represent the different types of comparisons recommended by the AMS workshop. To provide a common basis for the comparison of model predictions and data, it was necessary to account for background concentration. The background for the Oklahoma data sets was 3.0 parts/10^{15} and for the SRP data sets, 16.0 pCi/m^3. The main assumptions made in the creation of the statistical data sets are described below.

(a) The Oklahoma and SRP model/data comparisons are evaluated separately in terms of the AMS statistics.

(b) The Oklahoma data sets have averaging periods of 45 minutes (100 km arc) and 3 hours (600 km arc). Model/data comparisons for both averaging periods at Oklahoma were not separated for two reasons. First, it is expected that the usual damping due to time averaging should not be very significant in comparing 3 hour averages and 45 minute averages[50]. Second, if a split in model/data pairs were made, there would not be a sufficient number of points in either set to carry out meaningful statistics.

(c) A similar issue to (b) arises for the Savannah River Plant data sets. Most of the observed values represent 10 hour averaging periods. For roughly 10% of all points, the averages were over a different period with 14 hours the most common alternative. These different averaging periods were included together with the 10-hour averaging periods in the statistical data sets since the numbers of points involved and the differences in time periods were not significant.

Table 4-1. Summary of Data Sets for Use with AMS Statistics (Examples Given for SRP Data Base Only).

A. Peak Concentration

(A-1) Compare highest observed value for each event with highest prediction for same event (paired in time, not location).

(A-2) Compare highest observed value over all 65 experimental periods at each monitoring station with the highest prediction for those 65 experimental periods at the same station (paired in location, not time).

(A-3a) Compare maximum observed value for the 65 experimental periods with highest predicted values representing different time or space pairing (fully unpaired; paired in location; paired in time; paired in space and time).

(A-3b) Compare maximum predicted value for the 65 experimental periods with highest observed values for various pairings, as in (A-3a).

(A-4a) Compare highest N (=25) observed and highest N predicted values, regardless of time or location.

(A-4b) Compare highest N (=25) observed and highest N predicted values, regardless of time, for a given monitoring location. (A total of 13 data sets.)

(A-5) Same as (A-4a), but for subsets of events by meteorological conditions (stability and wind speed).

B. All-Concentrations

(B-1) Compare observed and predicted values at a given station, paired in time (a total of 13 data sets).

(B-2) Compare observed and predicted values for a given time period, paired in space (not appropriate for data sets with few monitoring sites).

(B-3) Compare observed and predicted values at all stations, paired in time and location (one data set) and by time of day.

(B-4) Same as (B-3), but for subsets of events by meteorological conditions (stability and wind speed) and by time of day.

(d) In order not to dilute the statistics with numerous points that had both the predicted value at background and the observed value at or below background, such pairs were discarded from all statistical (and graphical) data sets.

In this application of the AMS statistics, no outliers were defined and discarded. This treatment is consistent with the application of the AMS statistics to the EPA evaluation of rural, urban, and complex terrain models. The significance of outliers was considered to be less important due to the often counterbalancing effect of allowing both (a) large predicted values with zero observed, and (b) large observed values with zero predicted. The elimination of any outliers would be inappropriate, in any case, for the category A statistical data sets (see Table 4-1) since these data sets relate to an evaluation of the prediction of peak concentrations.

4.2.3. Peak Concentrations

For peak concentrations, comparisons are made to determine model performance both on an unpaired basis and for various pairings in time and space. The first two items in Table 4-1 represent a comparison of the highest observed and highest predicted concentrations paired in time (A-1) and paired in location (A-2).

The (A-3a) comparison includes the highest observed concentrations, regardless of time and space, and predicted values representing different time and space pairings. Item (A-3b) is directly analogous to (A-3a), but starts from the highest predicted value. Data sets (A-3a) and (A-3b) were not examined as part of this evaluation because it was felt that their inclusion would not provide additional useful information based on the nature of the database. The data sets (A-3a) and (A-3b) were dropped from further consideration.

Items (A-4) and (A-5) involve comparisons of the "N" highest observed and predicted values, unpaired in time or space. The AMS workshop recommended that such comparisons be based on the upper 2 to 5 percent of concentrations, rather than on one or two extreme values. In the previous three evaluations (rural, urban, and complex terrain), this percentage was replaced by a small number (N=25) of points. This number was expected to be appropriate in representing the set of highest observed and predicted values, and at the same time in still providing a statistical basis for the determination of confidence limits. In the present evaluation, the minimum of either 25 or 25% of the observed or predicted values is used instead. In preparing statistical calculations, any statistical data set with fewer than 10 elements was not considered further.

One difference between the present and past use of the AMS statistics by EPA is the neglect in this study of spatial and temporal correlation that may exist in the data, especially over a period of a few hours. In the complex terrain evaluations[50], employing 1-hour and 3-hour periods separately, the highest 25 values were screened to eliminate cases with two or more high values from the same period, or with two consecutive high values at the same location. That screening was intended to reduce the effects of autocorrelation and to avoid double-counting a single event. Although some correlations were anticipated in the Oklahoma and SRP data, it was expected that auto-correlation corrections would be small. Such corrections[57] were, therefore, not applied. Moreover, it was not expected that serial correlation would be a serious enough problem in the Oklahoma and SRP cases to lead to a significant impact on the results. The probabalistic statements often made on the basis of the AMS statistics have much greater uncertainty than any correction autocorrelation would provide.

Comparisons of the highest 25 observed and predicted values were performed for all stations combined (A-4a), but not for each station individually or for selected meteorological conditions. Further separation of the data points into meaningful subcategories would produce data sets that would be too small to provide meaningful statistics.

4.2.4. Comparisons of All Concentrations

As requested by the AMS workshop, comparisons were also made based upon all observed and predicted concentration values. Item (B-1) is the comparison of observed and predicted values at a given monitoring station (for all data pairs where the predicted or observed concentration was above background). Item (B-3) represents comparisons based on the set of values from all stations combined. Item (B-4) was omitted as it represents subsets of (B-3) that do not contain enough points for meaningful statistics.

4.2.5. Statistical Analysis of Model Performance

The AMS workshop report recommended two somewhat different lists of performance measures for comparing model predictions with air quality data, one appropriate for data sets representing pairs of observed and predicted values (Table 4-2), the other appropriate for unpaired data sets (Table 4-3). Paired data sets provide a means for assessing how well a model predicts on an event-by-event basis, whereas unpaired sets do not. Table 4-4 summarizes the basic list of performance measures, and the statistical methods recommended for establishing confidence limits on each measure. At the head of each column (paired and unpaired) are listed the data sets from Table 4-4 to which each list of measures and statistical indicators has been applied in the present study.

The data sets from item (A-1) (highest observed and predicted values for each event) and from items (B-1), (B-3), and (B-4) all represent observed and predicted values paired in time. For these sets, statistical analyses based on the residual (i.e., the differences between each pair of observed and predicted values) are appropriate for measuring model performance. If the time pairing for these data sets is ignored, however, it is also possible to assess model performance (in aggregate) by comparing the features of the composite set of all observed values to those of the predicted values. Consequently, both paired and unpaired comparisons were recommended by the AMS workshop for these data sets. Data sets representing comparisons of the highest 25 values, regardless of time and space, provide no basis for paired analysis. For these sets ((A-4), (A-5)), only unpaired comparisons are to be performed. Item (A-2) represents a comparison of the single highest observed and predicted values from each of the N stations. Only the paired comparison performance measures were computed for this case. No statistics were computed for the single-value comparisons in item (A-3).

For paired comparisons, as noted above, the performance measures are based on an analysis of residuals. Model bias is indicated by the average and/or the median residual, with a value of zero representing no bias. The

Table 4-2. Statistics Recommended by AMS for Application to Data Sets Representing All-Concentration Comparisons.

	Highest per Event Paired in Time (A-1)	Highest per Station Paired by Location (A-2)	All Events Paired in Time and Location (B-3)	All Events at Each Station Paired in Time (B-1)	Subsets of Events Paired in Time and Location (B-4)
Number of events	A	A	A	A	A
Average observed	A	A	A	A	A
Average difference	A(CI)	A(CI)	A(CI)	A	A
Fraction O > P	x	A	A	x	x
Characteristic discrepancies					
σ_d	A(CI)	A(CI)	A(CI)	A	A
RMSE	x	A	A	x	x
AAR	x	A	A	x	x
Correlation coefficients					
Pearson R	x	A	A	x	x
Spearman ρ	x	A	A	x	x
Kendall τ	x	A	A	x	x
Variance comparison	x	A	A	x	x
Maximum frequency difference	A(CI)	A(CI)	A(CI)	x	x

A = test is applicable
A(CI) = test is applicable with confidence interval
x = test is not applicable
O = observed concentration
P = predicted concentration
σ_d = standard deviation of bias
RMSE = root mean square error
AAR = average absolute residual

Table 4-3. Statistics Recommended by AMS for Application to Data Sets Representing Peak Values Unpaired (25 Highest).

	Average Observed	Average Predicted	Difference of Averages	Difference of Medians	Variance Ratio	Freq. Dist. Comparison
All stations/ all events (A-4a)	A	A	A(CI)	A(CI)	A(CI)	A(CI)
By station/ all events (A-4b)	A	A	A	x	x	x
Subsets by met. conditions (A-5)	A	A	A	x	x	x

A = test is applicable
A(CI) = test is applicable with confidence interval
x = test is not applicable

Table 4-4. Statistical Estimators and Basis for Confidence Limits for General Performance Measures Recommended by AMS.

Performance Measure	Estimator	Basis for Confidence Interval Paired Comparison e.g. (B-1), (B-3)	Unpaired Comparison e.g. (A-1), (A-4a)
Bias	Average	One sample "t," with adjustment for serial correlation	Two sample "t"
	Median	Wilcoxon matched pair	Mann-Whitney
Noise/Scatter	Standard deviation of residues	Chi-square test on standard deviation of residuals	F test on variance ratio
	Gross variability	None	Not applicable
	Average absolute residual	None	Not applicable
Correlation	Pearson correlation coefficient	None	Not applicable
	Spearman's rank-correlation coefficient, ρ	None	Not applicable
	Kendall's τ	None	Not applicable
Frequency distribution comparison	Maximum difference between two cumulative distribution functions	Not applicable	Kolmogorov-Smirnov (K-S) test on $f_{obs.}$ versus $f_{pred.}$

characteristic magnitude of the residuals is an indicator of the scatter between observed and predicted values on an event-by-event basis. Three measures of noise or scatter were computed:

- * Variance $\frac{1}{N-1} \sum_i (d_i - \bar{d})^2$
- * Gross variability $\frac{1}{N} \sum_i d_i^2$
- * Average absolute residual $\frac{1}{N} \sum_i |d_i|$

where d_i is the residual (observed (O) minus predicted (P)) for data pair i, $\bar{d}$ is the average residual, and N is the number of data pairs. The correlation of paired observed and predicted values is measured by the Pearson correlation coefficient.

For unpaired comparisons, the list of performance measures is shorter. Model bias is indicated by the difference between the average (or median) observed value and the average (or median) predicted value. A ratio of the variances of the observed and predicted values is provided to indicate whether the distribution of values in the two data sets is comparable. Similarly, the frequency distribution of observed values is compared with the frequency distribution of predicted values.

Standard statistical methods have been used to estimate confidence limits for each of the performance measures. For paired comparisons, the confidence interval on the average residual was estimated using a one-sample t test. This parametric test incorporates the assumption that the residuals follow a normal distribution; however, for large N, departures from normality are not critical. Under certain circumstances, serial correlation can affect results significantly. The AMS workshop recommended the adjustment of confidence limits for serial correlation. As noted earlier, the correction for serial correlation has been ignored since (a) serial correlation is not expected to be a significant problem with the Oklahoma and SRP model/data comparisons, and (b) the correction for serial correlation is not a complete one in any case.

A nonparametric indicator of model bias, analogous to the t test, is the median residual. The statistical method of estimating a confidence interval on the median residual is provided by the Wilcoxon matched-pairs test.

A confidence interval for the variance of the residuals is calculated using a chi-square test. No adjustment was made for serial correlation. No standard method is available for estimating confidence intervals for the gross variability or average absolute residual measures.

Comparison of two cumulative distribution functions is accomplished using the Kolmogorov-Smirnov (K-S) test. For this test, the two distribution functions are compared across the full range of concentration (or residual) values, and the maximum frequency difference between the two functions is identified.

For unpaired comparisons, two bias measures are computed. The average of the observed values is compared with the average of the predicted values. The confidence interval on the difference of the averages is estimated with a two sample t test. The median difference is also computed, and the confidence interval is estimated using the Mann-Whitney nonparametric test.

The variance of observed values is compared with the variance of predicted values for unpaired data sets. The performance measure is the ratio of the variances; the F test provides confidence limits on the ratio. The frequency distribution comparison for unpaired data sets provides a measure of the difference between the observed and predicted distribution functions. The K-S test is again used to assess the statistical significance of the maximum frequency difference.

4.2.6. Model Performance Results using AMS Statistics

Statistics comparing observed and predicted concentrations have been generated for each of the eight short-term long-range transport models and the two data bases. The results for each data base are presented at the same time for each statistical data set.

(A-1) ... Statistics for Highest Concentration by Event

Statistics were prepared for the data set consisting of the highest observed and predicted concentrations over the entire monitoring network for each sampling period, paired in time. For this statistical data set, pairing is in time and not space. Tables 4-5 and 4-6 present the results for the

Table 4-5. Statistical Data Set (A-1) for Oklahoma Data ..... (parts per 10^{15}). (Compares highest observed value for each event with highest prediction for same event, paired in time, not location).

Model	Number of Events	Average Observed Value	Average Difference* (Obs-Pred)	Standard Deviation of Residuals*	Maximum Frequency Difference*
MESOPUFF	21	753.1	-437.9	1023.6	0.29
			(-903.8, 28.1)	(783.1, 1478.2)	(0.420)
MESOPLUME	21	753.1	-836.1	2181.4	0.24
			(-1829.1, 156.9)	(1668.9, 3150.0)	(0.420)
MSPUFF	21	753.1	-3130.1	4955.8	0.33
			(-5386.1, -874.2)	(3791.5, 7156.5)	(0.420)
MESOPUFF II	21	753.1	-363.9	1171.0	0.29
			(-896.9, 169.2)	(895.9, 1691.0)	(0.420)
ARRPA	21	753.1	-827.2	2024.0	0.33
			(-1748.5, 94.2)	(1548.5, 2922.8)	(0.420)
RADM	21	753.1	-1756.1	2740.5	0.33
			(-3003.6, -508.6)	(2096.6, 3957.4)	(0.420)
RTM-II	21	753.1	404.6	1322.5	0.24
			(-197.4, 1006.6)	(1011.8, 1909.8)	(0.420)

* 95 percent confidence in parenthesis.

Table 4-6. Statistical Data Set (A-1) for Savannah River Plant Data (pCi/m^3). (Compares highest observed value for each event with highest prediction for same event, paired in time, not location).

Model	Number of Events	Average Observed Value	Average Difference* (Obs-Pred)	Standard Deviation of Residuals*	Maximum Frequency Difference*
MESOPUFF	63	228.4	-104.0	903.5	0.30
			(-331.5, 123.6)	(768.3, 1096.5)	(0.242)
MESOPLUME	63	228.4	-87.5	744.5	0.37
			(-275.0, 100.0)	(633.1, 903.6)	(0.242)
MSPUFF	63	228.4	-54.2	616.9	0.46
			(-209.6, 101.1)	(524.6, 748.7)	(0.242)
MESOPUFF II	63	228.4	-26.1	644.3	0.21
			(-188.3, 136.2)	(547.9, 782.0)	(0.242)
MTDDIS	63	228.4	-168.0	531.2	0.27
			(-301.8, -34.2)	(451.7, 644.6)	(0.242)
RADM	63	228.4	-247.8	853.0	0.29
			(-462.6, -32.9)	(725.3, 1035.2)	(0.242)
RTM-II	63	228.4	-27.3	489.6	0.19
			(-150.6, 96.0)	(416.4, 594.2)	(0.242)

Oklahoma and SRP data sets respectively. The number of events for Oklahoma is 21 which includes the events on the 100 km and 600 km arcs and data from both experiments. The number of events for the Savannah River Plant data base is 63. Two events were dropped from consideration because there were fewer than three stations reporting measured concentrations above background for those events. The formula for the 95% confidence interval for the maximum frequency difference is a function of the number of cases only, $1.36\sqrt{(2/N)}$. The value, therefore, is the same (0.420 or 0.242) for all models since the number of cases is always 63 or 21.

The average differences displayed in both tables indicate that each of the models for each data base tends to overpredict. The only exception is the RTM-II model at Oklahoma. For that data set, the large spreading observed with the RTM-II model (with low peak values) led to the inclusion of more predicted/observed pairs with lower predicted values. The largest overprediction occurs for MSPUFF for Oklahoma and for RADM for the SRP data base. RADM is also the only model that overpredicts at the 95% confidence level for both data sets (based on the number of samples available).

The standard deviation of residuals is an indicator of the range of residual values encountered for each model. The smallest standard deviation was observed from MESOPUFF at Oklahoma and RTM-II at the Savannah River Plant. The largest occurred with MSPUFF at Oklahoma and MESOPUFF at the Savannah River Plant.

Comparisons of frequency distributions of observed and predicted concentrations ignore any time pairing between predicted and observed values. The cumulative distribution function f(C) represents the fraction of the data set (in this case, the fraction of either 21 or 63 data points) with concentration values less than or equal to C. The value presented in this column is the largest absolute difference between the observed and predicted distribution functions (for the same concentration value) when the two functions are compared over all concentration values. The value given in parentheses is the maximum difference which is significantly different from zero, at a 95% confidence level, as given by the Kolmogorov-Smirnov (K-S) test. This confidence interval is a function of the number of cases as presented above.

There are two ways to interpret the results of the statistical estimator entitled "maximum frequency difference." First, one may say that the difference between the frequency distribution of predicted concentrations and the frequency distribution of observed concentrations was not significantly different from zero for all the models at Oklahoma and only for MESOPUFF II and RTM-II for the SRP data base. Second, this may be interpreted as indicating that for all the models at Oklahoma and for MESOPUFF II and RTM-II at SRP, the predicted and observed values in this statistical data set come from the same distribution at the 95% confidence level.

(A-2) ... Statistics on Highest Concentration at each Station

Tables 4-7 and 4-8 present performance statistics which compare the maximum observed and predicted concentration values at each monitoring station. These statistics involve a different number of stations for each model at Oklahoma and SRP (dependent on the location and spreading patterns of the models). The fact that at Oklahoma there are only 28 stations available for RADM, and 37 and 38 for RTM-II and MSPUFF, respectively, is indicative of the differences in spreading characterisics and correct positioning of the predicted plumes simulated by the models.

The statistics provided in these two tables compare observed and predicted values for the number of data pairs shown in the first column. The next two columns present the average of the observed concentrations and the average differences between observed and predicted values. The 95 percent confidence interval is given in parentheses, as calculated with a one-sample t test. For the SRP data base, this average difference indicates overprediction for all models; the overprediction is significant at the 95% confidence level for MESOPUFF and MESOPLUME. For the Oklahoma data base, overprediction occurs for four of the seven models tested (MESOPLUME, MSPUFF, RADM, and ARRPA). In fact, for RTM-II, the model shows a significant underprediction at the 95% confidence level.

The fourth column in these tables displays the fraction of positive residuals. This performance measure indicates the fraction of predicted/observed data pairs for which the observed concentration is larger than the predicted concentration. The results indicate underprediction at Oklahoma for all models except MSPUFF (50%) and ARRPA (42%). The ARRPA results indicate overprediction. It is interesting that at Oklahoma, the two measures of bias (average difference and fraction of positive residuals) lead to generally

Table 4-7. Statistical Data Set (A-2) for Oklahoma Data (parts per 10^{15}). (Compares highest observed value over all 21 experimental periods at each monitoring station with the highest prediction for those 21 experimental periods at the same station, paired in location, not time).

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE DIFFERENCE* (OBS-PRED)	FRACTION OF POSITIVE RESIDUALS (OBS > PRED)	STANDARD DEVIATION OF RESIDUALS*	ROOT MEAN SQUARE ERROR	AVERAGE ABSOLUTE RESIDUAL	PEARSON CORR. COEFF.	SPEARMAN CORR. COEFF.	KENDALL CORR. COEFF.	VARIANCE COMPARISON* (OBS/PRED)
MESOPUFF	35	828.4	192.9 (-16.0, 401.9)	0.69	608.1 (491.3, 797.7)	629.7	393.7	0.91	0.17	0.10	1.171 (0.591, 2.321)
MESOPLUME	34	852.8	-143.3 (-612.0, 325.5)	0.56	1343.1 (1082.1, 1770.2)	1330.9	667.7	0.81	0.27	0.17	0.440 (0.220, 0.881)
MSPUFF	38	763.0	-887.4 (-1990.1, 215.4)	0.50	3354.3 (2731.9, 4344.1)	3426.8	1268.3	0.86	0.24	0.15	0.101 (0.053, 0.195)
MESOPUFF II	31	935.3	166.1 (-553.3, 885.5)	0.61	1961.2 (1567.2, 2621.5)	1936.4	1153.6	0.07	0.30	0.19	1.329 (0.641, 2.756)
RTM-II	37	783.6	433.1 (64.0, 802.3)	0.54	1107.0 (899.3, 1439.5)	1174.7	599.1	0.71	0.39	0.28	5.985 (3.081, 11.624)
RADM	28	1035.5	-1145.6 (-2862.4, 571.2)	0.64	4427.5 (3500.7, 6027.2)	4496.1	2368.6	-.05	0.32	0.19	0.153 (0.071, 0.330)
ARRPA	33	878.6	-829.1 (-1620.3, -37.9)	0.42	2230.8 (1791.8, 2954.7)	2348.0	1466.2	0.57	0.37	0.28	0.305 (0.151, 0.618)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

Table 4-8. Statistical Data Set (A-2) for Savannah River Plant Data (pCi/m^3). (Compares highest observed value over all 65 experimental periods at each monitoring station with the highest prediction for those 65 experimental periods at the same station, paired in location, not time).

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE DIFFERENCE* (OBS-PRED)	FRACTION OF POSITIVE RESIDUALS (OBS > PRED)	STANDARD DEVIATION OF RESIDUALS*	ROOT MEAN SQUARE ERROR	AVERAGE ABSOLUTE RESIDUAL	PEARSON CORR. COEFF.	SPEARMAN CORR. COEFF.	KENDALL CORR. COEFF.	VARIANCE COMPARISON* (OBS/PRED)
MESOPUFF	13	621.3	-596.2 (-1180.6, -11.9)	0.15	967.0 (693.4, 1596.2)	1103.9	724.8	0.80	0.76	0.56	0.262 (0.080, 0.862)
MESOPLUME	13	621.3	-409.2 (-778.8, -39.6)	0.15	611.6 (438.5, 1009.6)	716.1	577.2	0.76	0.61	0.49	0.612 (0.186, 2.011)
MSPUFF	13	621.3	-470.7 (-1023.0, 81.5)	0.15	913.9 (655.3, 1508.6)	996.3	640.2	0.60	0.76	0.62	0.419 (0.128, 1.378)
MESOPUFF II	13	621.3	-123.7 (-532.2, 284.8)	0.23	676.0 (484.7, 1115.9)	661.2	492.5	0.59	0.77	0.62	0.974 (0.296, 3.199)
MTDDIS	13	621.3	-161.7 (-558.4, 235.1)	0.31	656.6 (470.8, 1083.8)	651.2	443.8	0.66	0.70	0.51	0.795 (0.242, 2.612)
RTM-II	13	621.3	-48.9 (-441.7, 343.8)	0.23	650.0 (466.1, 1072.9)	626.4	507.0	0.51	0.64	0.49	2.156 (0.656, 7.084)
RADM	13	621.3	-3955.6 (-11465.2,3554.1)	0.31	12427.2 (8910.7,20513.5)	12577.8	4152.3	0.25	0.70	0.53	0.003 (0.001, 0.011)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

contradictory findings. At SRP, both measures of bias (average difference and fraction of positive residuals) indicate overprediction.

The next three performance measures provide estimates of scatter. They include the standard deviation of residuals (with 95% confidence limits calculated from the F test), root mean square error, and average absolute residual. MESOPUFF has the smallest values for Oklahoma, and RTM-II, MESOPUFF II and MTDDIS have the smallest values for SRP. RADM has the largest values at both sites. A single large outlier predicted by RADM at SRP may, to a large degree, be causing the very large scatter indicated by these statistics.

The Pearson correlation coefficient of observed and predicted concentration pairs and the nonparametric Spearman correlation of ranked sets of observed and predicted concentrations provide indications of the spatial correlation of the maximum concentration values at each station. Pearson's coefficient varied greatly at Oklahoma (-0.05 to 0.91) and at SRP (0.25 to 0.8). The single large outlier for RADM in the SRP data base is the cause of the poor Pearson's coefficient at SRP; this outlier is effectively removed upon consideration of the nonparametric correlation coefficients, Spearman's ρ and Kendall's τ.

The last column in Tables 4-7 and 4-8 (variance comparison) presents the ratio of observed variance divided by the predicted variance, with 95 percent confidence bounds in parentheses as calculated by an F test. The variance comparisons for the highest predicted and observed concentrations by station generally reveal values less than unity for all models except RTM-II, which has variance comparison ratios greater than unity for both the Oklahoma and SRP sites. (For the Oklahoma site both MESOPUFF and MESOPUFF II exhibit values slightly above unity.) Values of variance comparison and confidence intervals less than unity indicate the large magnitude and range of predicted values compared to the observed values.

(A-4a) ... Statistics for Highest 25 Values

Statistics on the set of 25 highest observed and 25 highest predicted concentrations are presented for each model in Tables 4-9 and 4-10. The first two columns of results are simply the average of the 25 highest predicted values for each data set. Comparing the average observed and average

Table 4-9. Statistical Data Set (A-4a) for Oklahoma Data (parts/10^{15}).
(Compares highest N (=25) observed and highest N predicted values, regardless of time or location.)

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE PREDICTED VALUE	DIFFERENCE OF AVERAGES* (OBS-PRED)	DIFFERENCE OF MEDIANS* (OBS-PRED)	VARIANCE COMPARISON* (OBS/PRED)	MAXIMUM FREQUENCY DIFFERENCE
MESOPUFF	25	1809.7	2617.6	-807.9 (-1548.9, -66.9)	-1448.5 (-1716.6, -221.2)	1.341 (0.591, 3.045)	0.48 (0.385)
MESOPLUME	25	1809.7	3231.5	-1421.8 (-2511.3, -332.3)	-1536.2 (-2235.5, -174.7)	0.367 (0.161, 0.832)	0.40 (0.385)
MSPUFF	25	1809.7	6347.0	-4537.3 (-6940.9,-2133.7)	-1680.3 (-5296.6, -989.7)	0.060 (0.027, 0.137)	0.60 (0.385)
MESOPUFF II	25	1809.7	2230.4	-420.7 (-1105.3, 263.8)	-680.6 (-1120.6, -35.2)	2.042 (0.899, 4.635)	0.44 (0.385)
RTM-II	25	1809.7	1234.3	575.4 (-9.3, 1160.0)	-67.4 (-269.6, 772.6)	27.887 (12.284,63.305)	0.40 (0.385)
RADM	25	1809.7	5894.0	-4084.3 (-5391.0,-2777.7)	-3447.6 (-4512.5,-2507.4)	0.232 (0.102, 0.527)	0.64 (0.385)
ARRPA	25	1809.7	3183.5	-1373.9 (-2498.5, -249.2)	-1358.9 (-1991.9, -232.9)	0.337 (0.149, 0.765)	0.40 (0.385)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

Table 4-10. Statistical Data Set (A-4a) for Savannah River Plant Data (pCi/m^3). (Compares highest N (=25) observed and highest N predicted values, regardless of time or location.)

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE PREDICTED VALUE	DIFFERENCE OF AVERAGES* (OBS-PRED)	DIFFERENCE OF MEDIANS* (OBS-PRED)	VARIANCE COMPARISON* (OBS/PRED)	MAXIMUM FREQUENCY DIFFERENCE
MESOPUFF	25	643.7	926.2	-282.5 (-778.0, 213.0)	-191.5 (-363.5, 77.4)	0.250 (0.110, 0.569)	0.24 (0.385)
MESOPLUME	25	643.7	909.3	-265.6 (-645.0, 113.9)	-278.9 (-437.6, -8.9)	0.504 (0.222, 1.144)	0.36 (0.385)
MSPUFF	25	643.7	858.7	-215.0 (-628.9, 198.9)	-139.0 (-281.8, 110.8)	0.398 (0.175, 0.903)	0.20 (0.385)
MESOPUFF II	25	643.7	700.3	-56.6 (-384.6, 271.3)	-90.6 (-201.9, 116.5)	0.814 (0.359, 1.848)	0.20 (0.385)
TDDIS	25	643.7	925.4	-281.7 (-636.9, 73.5)	-293.6 (-401.9, -46.1)	0.619 (0.273, 1.406)	0.48 (0.385)
RTM-II	25	643.7	666.8	-23.2 (-312.6, 266.3)	-171.8 (-245.6, 71.0)	1.357 (0.598, 3.080)	0.32 (0.385)
RADM	25	643.7	3043.8	-2400.1 (-6153.6, 1353.5)	-766.2 (-905.0, -150.5)	0.004 (0.002, 0.008)	0.48 (0.385)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

predicted value indicates that all models except RADM (both data sets) and MSPUFF (Oklahoma only) are within a factor of two of the high 25 observed data. The first performance measure, presented in column 3, is the difference between the two averages. Negative values exist for all models except RTM-II for Oklahoma; the negative values indicate overprediction. The confidence intervals being negative as well for MESOPUFF, MESOPLUME, MSPUFF, RADM, and ARRPA at Oklahoma indicate that at the 95% confidence level, these models overpredict with respect to this statistic. The confidence intervals were determined using the two-sample Student t test.

The second performance measure relating to average values is the median difference (313th highest value) between all 625 possible pairings of the 25 highest observed and predicted concentrations. The 95% confidence interval is determined with the nonparametric Mann-Whitney test. Results for the median difference are very similar to those for the difference of averages. Here, all models exhibit overprediction for both data sets. With all models for Oklahoma (except RTM-II), overprediction is significant at the 95% confidence level.

The third performance measure is the variance ratio. The variance of the 25 highest observed values was divided by the variance of the 25 highest predictions. The F test was used to calculate the 95 percent confidence levels for these comparisons. Results indicate that, at SRP, the scatter of the 25 highest predictions is much larger than the scatter of the 25 highest observations, except for the RTM-II predictions. For Oklahoma, the scatter of the 25 highest predictions is much larger than the scatter of the 25 highest observations for the MESOPLUME, MSPUFF, RADM and ARRPA models. However, the MESOPUFF, MESOPUFF II and RTM-II models exhibit more scatter in the 25 highest observations than in the 25 highest predictions. The RTM-II model is extreme in having less scatter of its predictions than the observations indicate.

The last performance measure presented in Tables 4-9 and 4-10 is the frequency distribution comparison. The confidence interval is a function of the number of cases. The value is, therefore, the same (0.385) for all models, since the number of cases is always 25. For Oklahoma and for all models, one can reject the hypothesis that the predicted and observed peak values came from the same distribution (at the 95% confidence level). At SRP, the opposite is true (except for RADM and MTDDIS); one cannot reject the hypothesis that the peak values came from the same distribution.

(B-3) ... Statistics for All Concentrations Paired in Space and Time

Tables 4-11 and 4-12 present the comparison of all observed and predicted concentration values paired in time and space for the Oklahoma and SRP data bases, respectively. Results for Oklahoma show that in terms of average differences, an overprediction exists for all models except RTM-II. MSPUFF and RADM show significant overpredictions at Oklahoma and the RADM model does so again at SRP. For the MESOPUFF, MESOPLUME, MSPUFF and RADM models at Oklahoma, and for the MTDDIS at SRP, the entire confidence interval for average differences is negative. This indicates overprediction at the 95% confidence level for those four models at Oklahoma and for MTDDIS at SRP.

Values for the fraction of positive residuals, the standard deviation of residuals, the root-mean-square error, and the average absolute residual all exceed the average observed values. The largest values for the measures of scatter at Oklahoma occur for the MSPUFF and RADM models. The RADM model has the largest values for the measures of scatter at SRP. Significant scatter exists for all models at the 95% confidence level.

Correlations of observed and predicted concentrations are extremely low and negative in many cases. Variance ratios indicate that variances for predicted values are generally much greater than variances for observed values. The only exception is RTM-II (for both data sets) with a variance comparison exceeding 1.0.

4.3. GRAPHICAL EVALUATION OF THE EIGHT MODELS WITH THE OKLAHOMA AND SAVANNAH RIVER PLANT DATA

Graphical comparisons of predicted and observed concentrations for each data base are used to provide insight into systematic differences between model predictions and data. Four useful types of graph are:

(a) scatter plots of observed and predicted concentrations for all 15 SRP data sets and (separately) both Oklahoma data sets. Points paired in space and time are used in this scatter plot. A related graph presents a scatter plot of residuals (observed minus predicted) versus the average of the predicted and observed value for that pair,

Table 4-11. Statistical Data Set (B-3) for Oklahoma Data (parts per 10^{15}). (Compares observed and predicted values at all stations, paired in time and location).

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE DIFFERENCE* (OBS-PRED)	FRACTION OF POSITIVE RESIDUALS (OBS > PRED)	STANDARD DEVIATION OF RESIDUALS*	ROOT MEAN SQUARE ERROR	AVERAGE ABSOLUTE RESIDUAL	PEARSON CORR. COEFF.	SPEARMAN CORR. COEFF.	KENDALL CORR. COEFF.	VARIANCE COMPARISON* (OBS/FRED)	MAXIMUM FREQUENCY DIFFERENCE
MESOPUFF	224	243.4	-116.0 (-204.7, -27.3)	0.54	673.9 (616.7, 742.9)	682.3	290.4	0.68	0.13	0.08	0.65 (0.50, 0.84)	0.09 (0.13)
MESOPLUME	216	252.4	-191.3 (-331.9, -50.6)	0.61	1049.5 (958.9,1159.1)	1064.4	389.2	0.57	0.07	0.04	0.33 (0.25, 0.43)	0.19 (0.13)
MSPUFF	230	237.1	-538.5 (-824.5, -252.6)	0.52	2201.5 (2016.8,2423.5)	2261.8	708.5	0.75	0.08	0.04	0.07 (0.05, 0.09)	0.15 (0.13)
MESOPUFF II	203	268.6	-97.0 (-259.7, 65.6)	0.59	1176.3 (1071.8,1303.4)	1177.4	587.1	-0.14	-0.20	-0.14	0.90 (0.68, 1.18)	0.16 (0.13)
RTM-II	228	239.1	7.4 (-94.6, 109.4)	0.41	781.6 (715.8, 860.8)	779.9	355.0	0.12	0.24	0.17	3.23 (2.49, 4.19)	0.18 (0.13)
RADM	192	234.0	-721.9 (-1067.2, -376.7)	0.61	2428.3 (2207.0,2699.0)	2527.3	1205.3	-0.11	-0.10	-0.06	0.12 (0.09, 0.16)	0.24 (0.14)
ARRPA	208	262.1	-207.5 (-417.8, 2.7)	0.58	1539.0 (1403.8,1703.1)	1549.3	640.2	-0.05	0.08	0.06	0.33 (0.25, 0.43)	0.16 (0.13)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

Table 4-12. Statistical Data Set (B-3) for Savannah River Plant Data (pCi/m^3). (Compares observed and predicted values at all stations, paired in time and location).

MODEL	NUMBER OF DATA PAIRS	AVERAGE OBSERVED VALUE	AVERAGE DIFFERENCE* (OBS-PRED)	FRACTION OF POSITIVE RESIDUALS (OBS > PRED)	STANDARD DEVIATION OF RESIDUALS*	ROOT MEAN SQUARE ERROR	AVERAGE ABSOLUTE RESIDUAL	PEARSON CORR. COEFF.	SPEARMAN CORR. COEFF.	KENDALL CORR. COEFF.	VARIANCE COMPARISON* (OBS/PRED)	MAXIMUM FREQUENCY DIFFERENCE
MESOPUFF	312	70.5	-12.4 (-61.7, 37.0)	0.74	442.6 (410.4, 480.4)	442.1	132.3	0.07	-0.08	-0.06	0.34 (0.27, 0.43)	0.39 (0.11)
MESOPLUME	307	71.6	-12.9 (-57.3, 31.4)	0.78	394.1 (365.1, 428.0)	393.6	133.7	0.05	-0.04	-0.04	0.50 (0.40, 0.62)	0.46 (0.11)
MSPUFF	337	65.2	-10.2 (-52.7, 32.2)	0.70	395.4 (367.6, 427.8)	395.0	129.4	-0.02	-0.30	-0.23	0.48 (0.38, 0.59)	0.33 (0.10)
MESOPUFF II	342	64.3	-8.9 (-44.0, 26.1)	0.53	329.2 (306.2, 355.9)	328.8	117.9	-0.01	0.01	0.00	0.83 (0.67, 1.03)	0.09 (0.10)
MTDDIS	452	48.6	-36.4 (-61.4, -11.4)	0.28	269.7 (253.1, 288.5)	271.8	86.9	0.35	0.43	0.31	0.53 (0.44, 0.64)	0.24 (0.09)
RTM-II	331	66.4	-4.5 (-36.8, 27.7)	0.63	297.7 (276.6, 322.3)	297.3	109.9	0.09	0.05	0.03	1.06 (0.86, 1.32)	0.20 (0.11)
RADM	294	74.8	-213.4 (-521.4, 94.7)	0.74	2681.4 (2430.5,2917.6)	2685.3	324.0	0.25	0.21	0.16	0.01 (0.01, 0.01)	0.43 (0.11)

* 95 PERCENT CONFIDENCE INTERVAL IN PARENTHESES.

(b) frequency histogram of residuals for each model and each data base. Points paired in space and time are used to determine residuals. This type of graph illustrates whether a balance exists between overprediction and underprediction,

(c) frequency histograms showing the distribution of predicted values and observed values for a model. The determination of whether the model predictions and observed data (unpaired) have roughly the same frequency of high and low values is of interest, and

(d) cumulative frequency distribution of predicted and observed concentrations over all data sets for the SRP and (separately) the Oklahoma data bases. Points paired in space and time are used to create the residuals here as well. This kind of graph is related to the Kolmogorov-Smirnov statistic (presented in Section 4.2) which is aimed at identifying whether the predicted and observed distributions are statistically the same.

A complete set of these graphs for both data bases and all eight models appears in Appendix E. A subset of those graphs are presented below which illustrate special or unique facets of the performance of the models. The full set of graphs in Appendix E support even more strongly the characteristics described below.

4.3.1. Scatter Plots

The scatter plots for all eight models are presented in Figures 4-1 to 4-7 for the Oklahoma data base. Each box represents one predicted/observed pair; the horizontal axis is the observed value in parts per 10^{15}, and the vertical axis is the predicted value in the same units. Perfect model/data agreement is represented by the dotted line. A point on the vertical axis has a nonzero predicted value, but a zero observed value. A point on the horizontal axis has a zero predicted concentration and a nonzero observed concentration. For the Savannah River Plant data base the scatter plots for the eight models are presented in Figures 4-8 to 4-14, where units are in pCi/m^3.

Each scatter plot has most of the predicted/observed pairs on the axes. Predicted values greater than zero (background) are often associated with

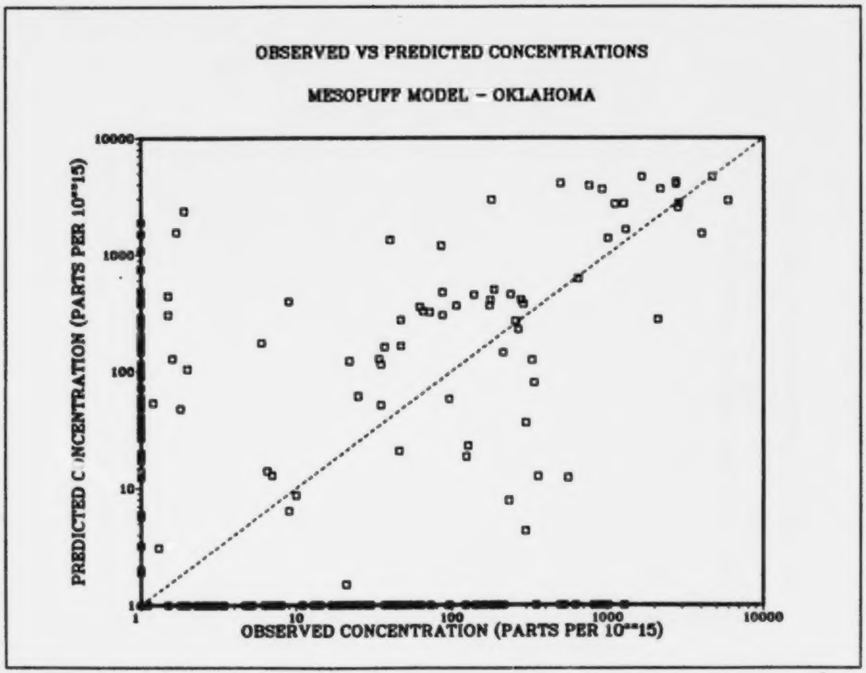


Figure 4-1. Scatter plot of observed and MESOPUFF predicted concentrations at Oklahoma (points paired in space and time).

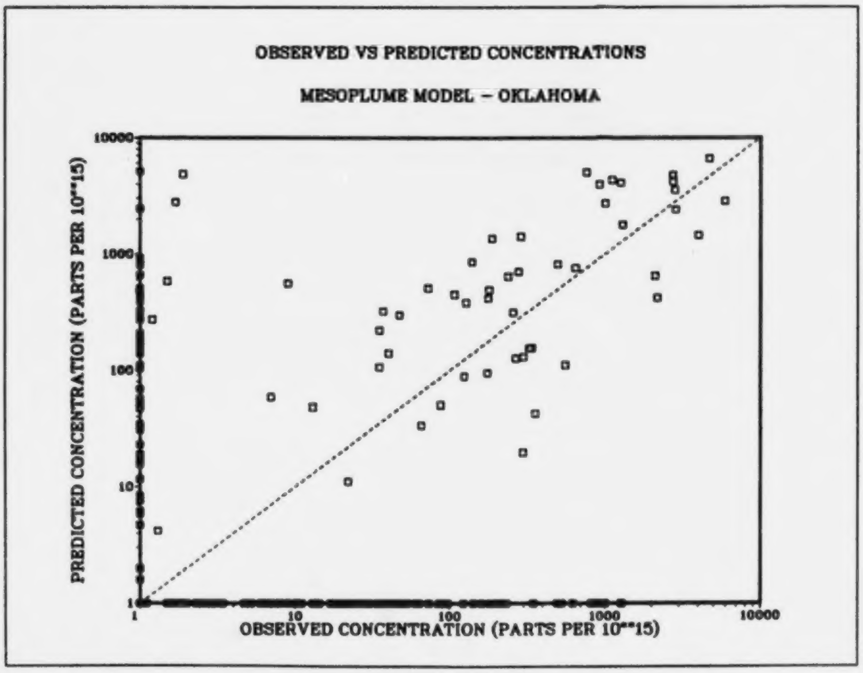


Figure 4-2. Scatter plot of observed and MESOPLUME predicted concentrations at Oklahoma (points paired in space and time).

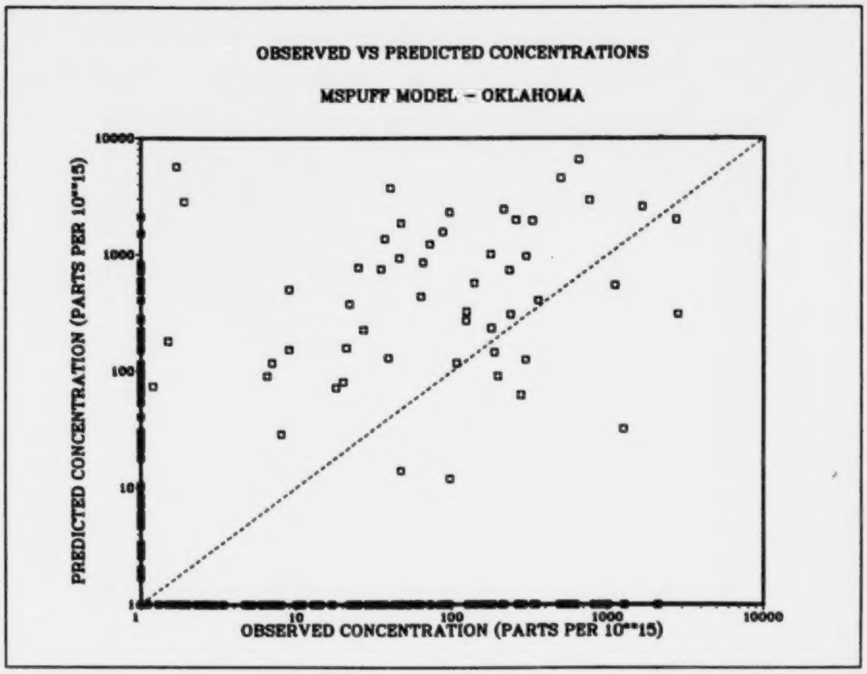


Figure 4-3. Scatter plot of observed and MSPUFF predicted concentrations at Oklahoma (points paired in space and time).

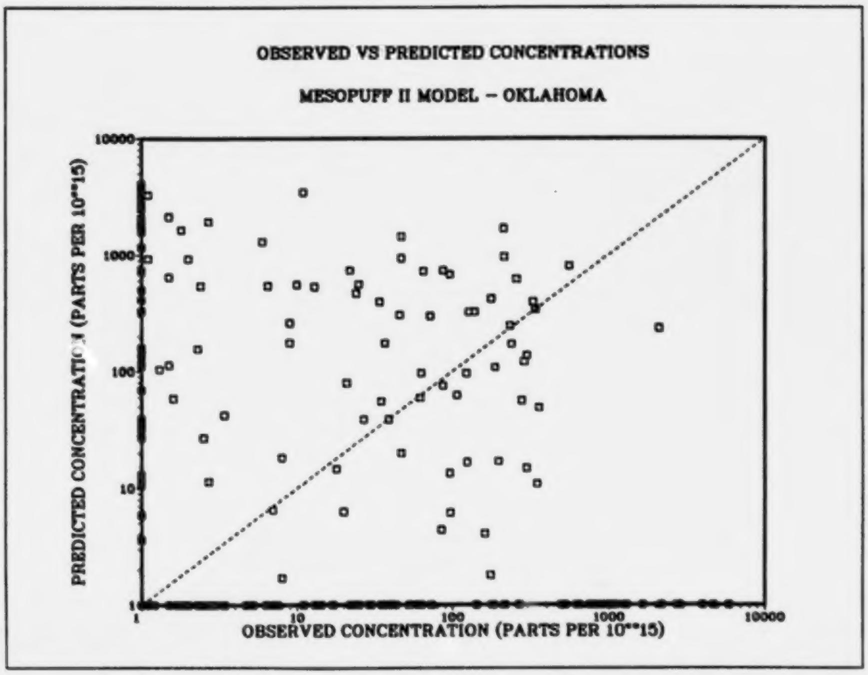


Figure 4-4. Scatter plot of observed and MESOPUFF II predicted concentrations at Oklahoma (points paired in space and time).

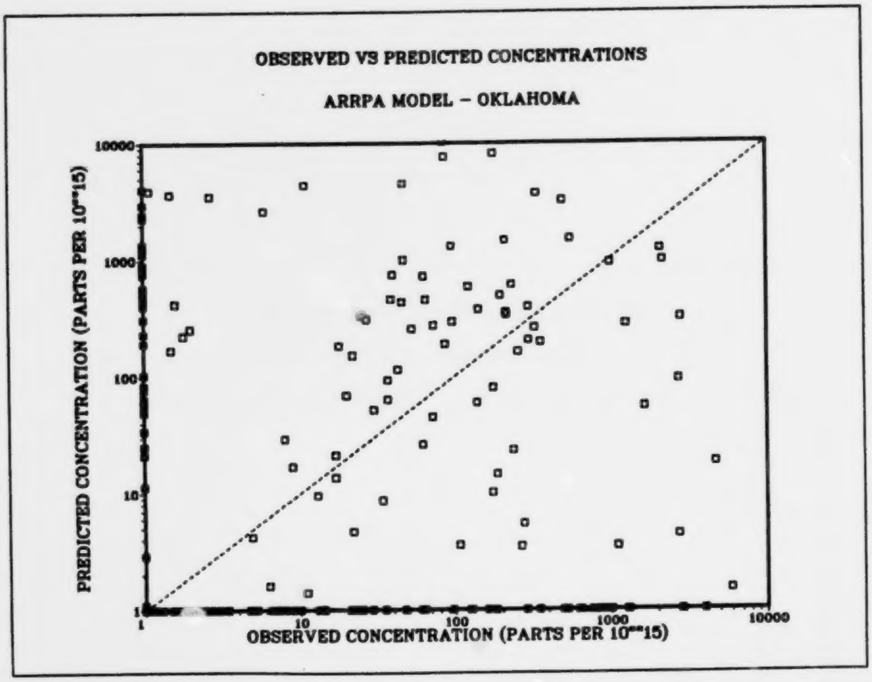


Figure 4-5. Scatter plot of observed and ARRPA predicted concentrations at Oklahoma (points paired in space and time).

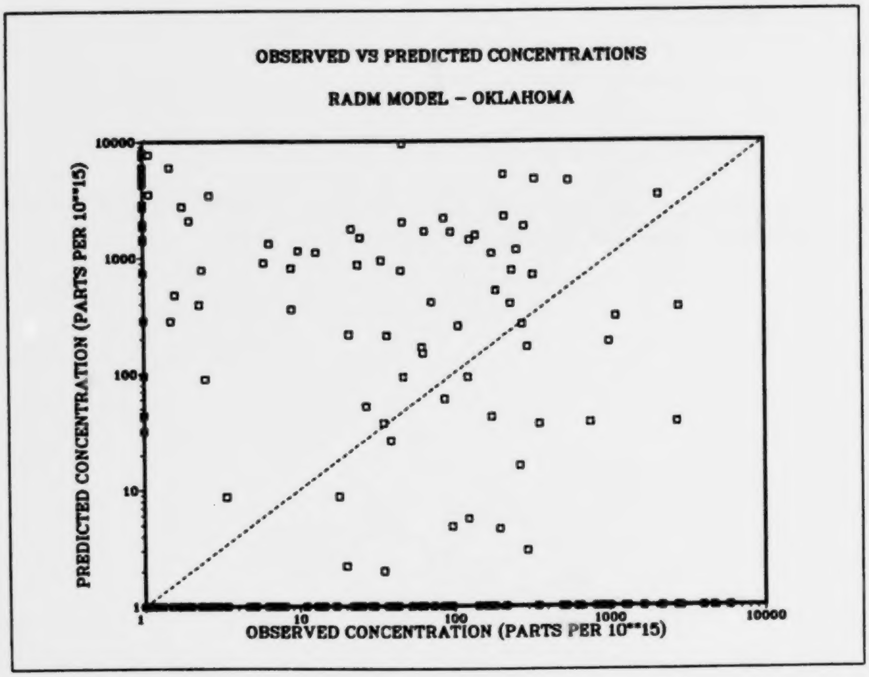


Figure 4-6. Scatter plot of observed and RADM predicted concentrations at Oklahoma (points paired in space and time).

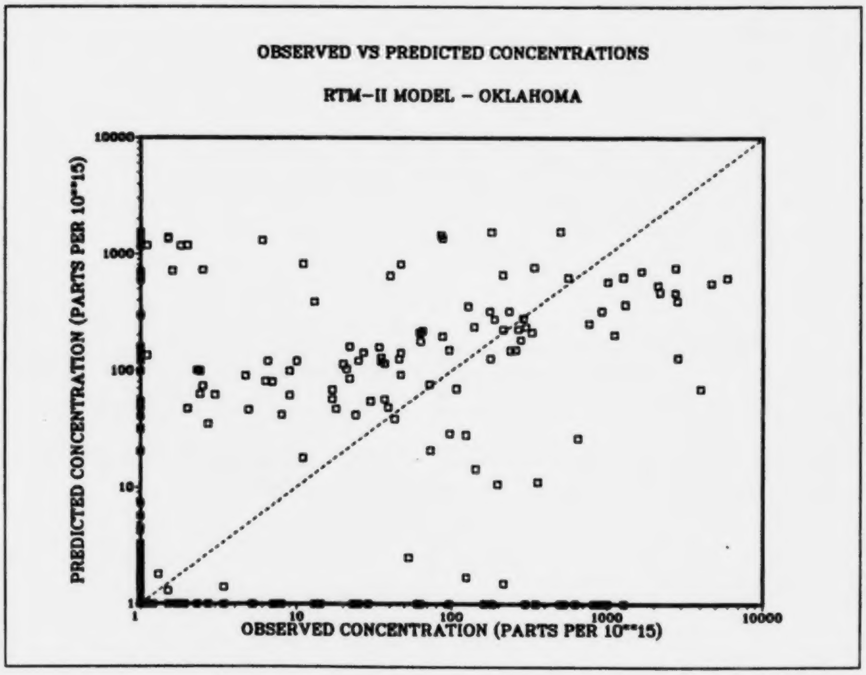


Figure 4-7. Scatter plot of observed and RTM-II predicted concentrations at Oklahoma (points paired in space and time).

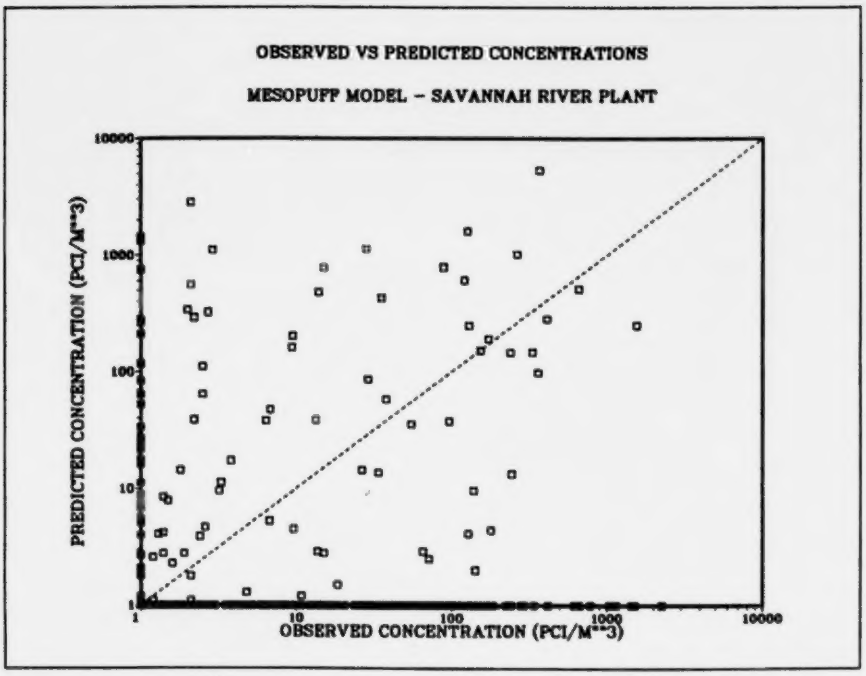


Figure 4-8. Scatter plot of observed and MESOPUFF predicted concentrations at Savannah River Plant (points paired in space and time).

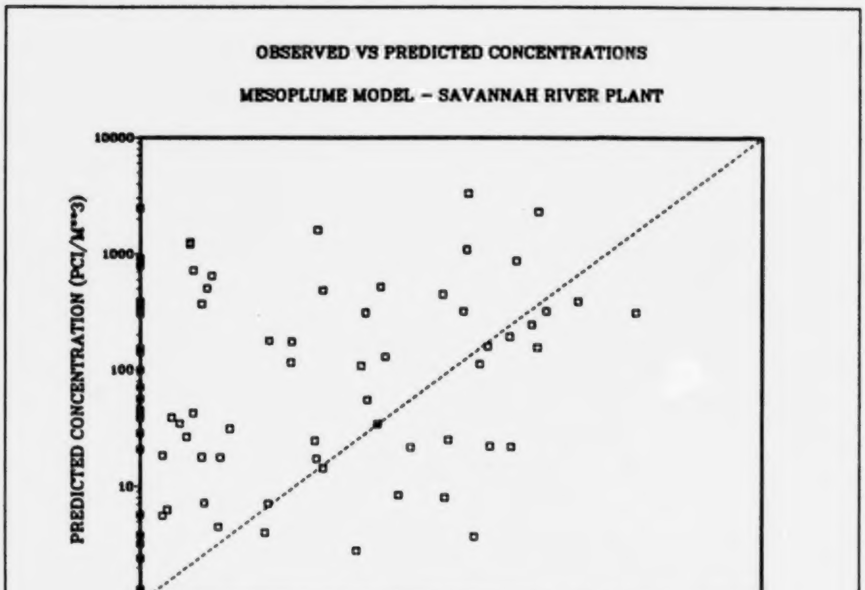


Figure 4-9. Scatter plot of observed and MESOPLUME predicted concentrations at Savannah River Plant (points paired in space and time).

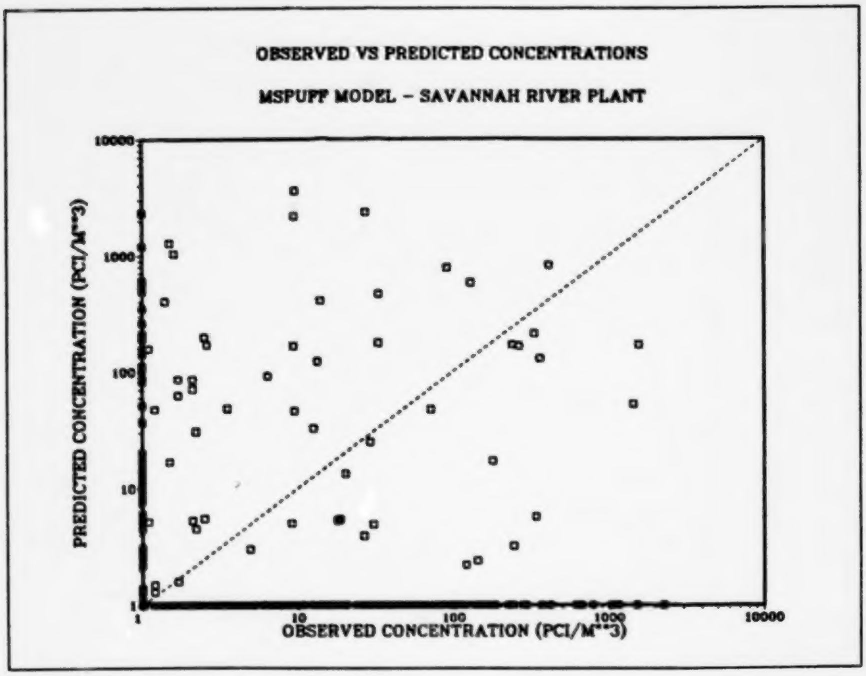


Figure 4-10. Scatter plot of observed and MSPUFF predicted concentrations at Savannah River Plant (points paired in space and time).

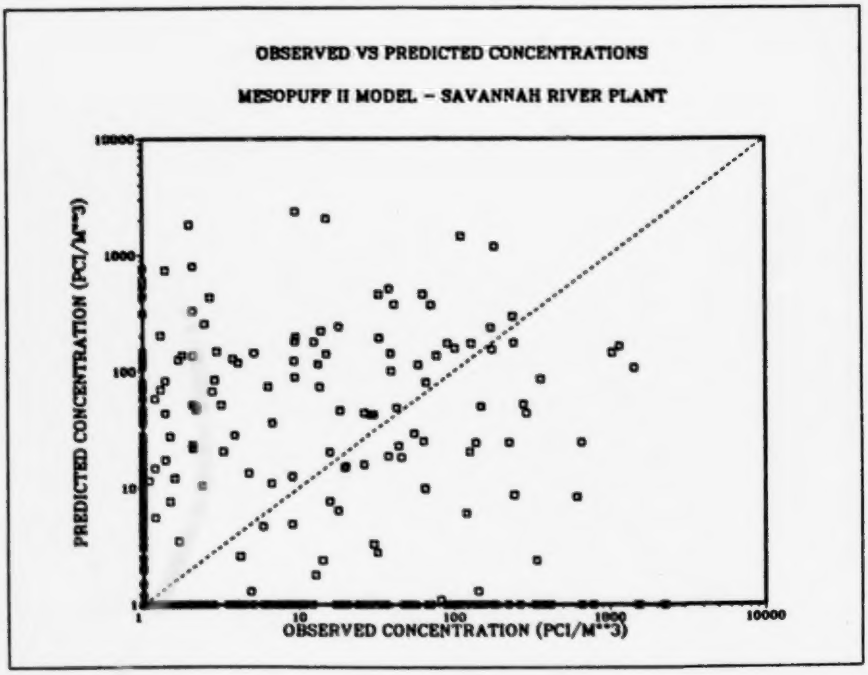


Figure 4-11. Scatter plot of observed and MESOPUFF II predicted concentrations at Savannah River Plant (points paired in space and time).

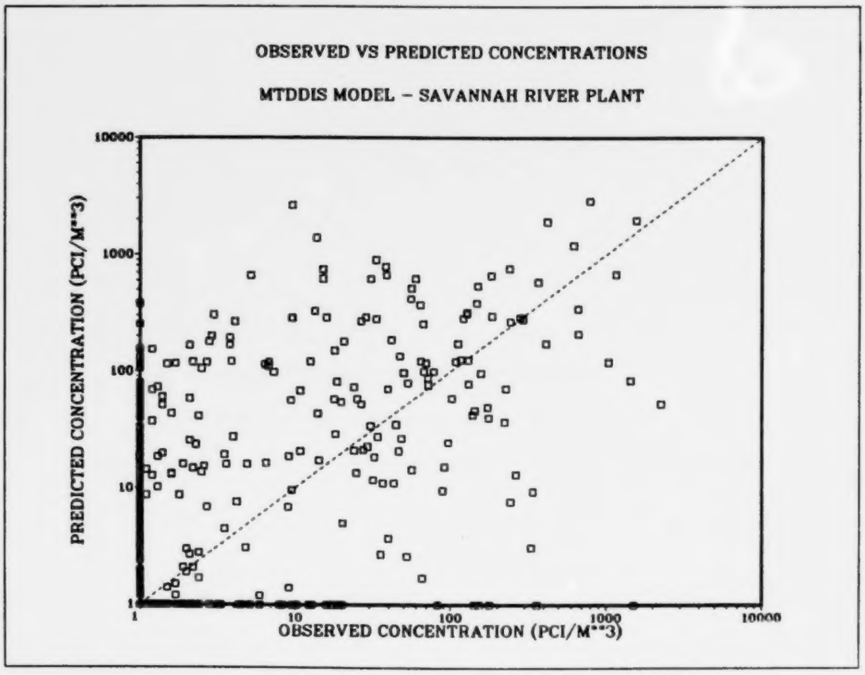


Figure 4-12. Scatter plot of observed and MTDDIS predicted concentrations at Savannah River Plant (points paired in space and time).

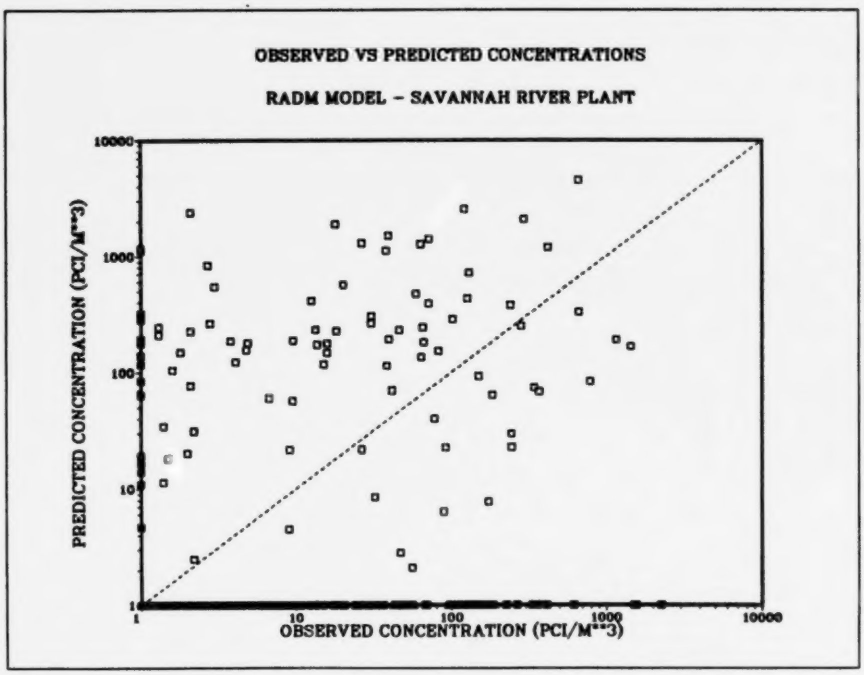


Figure 4-13. Scatter plot of observed and RADM predicted concentrations at Savannah River Plant (points paired in space and time).

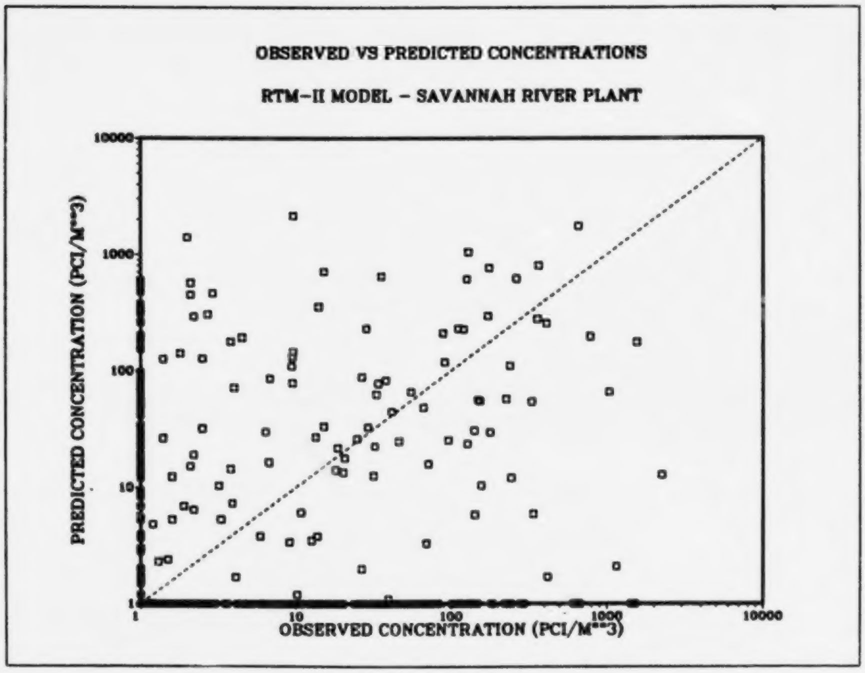


Figure 4-14. Scatter plot of observed and RTM-II predicted concentrations at Savannah River Plant (points paired in space and time).

observed values that are zero and vice versa. This is clear evidence that the predicted and observed plume patterns on the ground are offset from one another. Table 4-13 summarizes these behaviors.

Table 4-13. Comparison of Predictions of the Models Based on Percentage of Pairs where Observed Values are Greater than Predicted Values (points paired in space and time).

Model	Oklahoma	SRP
	(% of pairs where obs > pred)	
MESOPUFF	69	74
MESOPLUME	56	78
MSPUFF	50	70
MESOPUFF II	61	53
MTDDIS	*	28
ARRPA	54	*
RADM	64	63
RTM-II	42	74

* Model not applied to this data base.

As may be seen, most of the percentages are larger than 50% indicating that the plumes have generally not spread out over the observed data due either to (a) an underestimation of lateral spreading of the predicted plume, and/or (b) a poor location of the predicted plume with respect to the observed plume. The situation where a predicted plume underestimates spreading may be viewed as one where the observed plume is expanding wider than the predicted plume, leading to many points where the observed values are greater than zero and the predicted value are zero.

It is interesting to note that the RTM-II model for Oklahoma (with a 42% value) tends to overpredict plume spreading, whereas for SRP (with a 74% value) it tends to underpredict spreading. An examination of the complete set of contour plots (Appendix E) reveals this to be the case. The cause of this behavior is simply in the choice of the η-parameter (and its upper and lower bounds) in the lateral diffusivity for RTM-II. In this study predictions of horizontal spreading were found to be sensitive to the choice of these parameters. Apparently, too large a value of η and its lower bound were selected by the model developers for Oklahoma and too small a value of η and its upper bound were recommended by them for the SRP runs.

The other interesting result is for MTDDIS, which predicts that 28% of all predicted/observed pairs (points paired in space and time) have observations greater than predictions. One may infer from this comparison that the MTDDIS predicted plume overpredicts horizontal spreading. That conclusion will be supported further by the results of the pattern-comparison studies in Chapter 5.

The general tendency to underpredict spreading shown in Table 4-13 also coincides with the general tendency noted above of the models to overpredict peak values for the Oklahoma and SRP data sets.

4.3.2. Other graphical comparisons

The other types of graphical comparisons presented are the following:

(a) frequency distribution of predicted and observed concentrations, (like Fig. 4-15)
(b) cumulative frequency distributions of predicted and observed concentrations, (like Fig. 4-16)
(c) frequency distribution of residuals for points paired in space and time, (like Fig. 4-17)
(d) Average of observed and predicted concentrations versus residuals, (not discussed, like Fig. E-99)

Conclusions drawn from these four types of comparison are surprisingly similar among the models for a given data base. A short summary of trends follows:

4.3.2.1. Oklahoma Data Base

Table 4-14 provides a summary of the behaviors of the models for the Oklahoma graphical comparisons. In this Table, "P" represents predicted

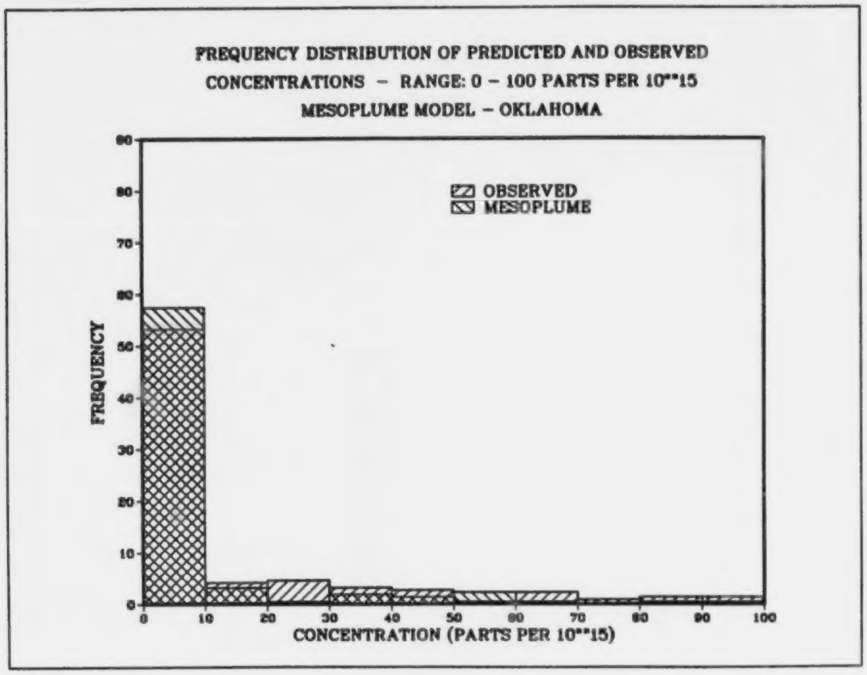


Figure 4-15. Frequency distribution of predicted and observed concentrations at Oklahoma for MESOPLUME based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}.

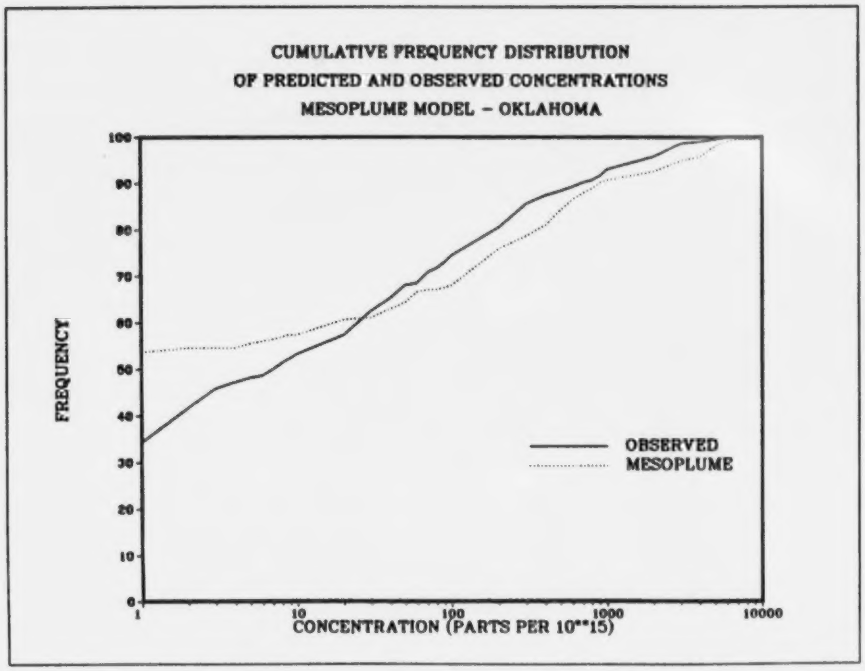


Figure 4-16. Cumulative frequency distributions of MESOPLUME predictions and observed concentrations at Oklahoma based on points paired in space and time.

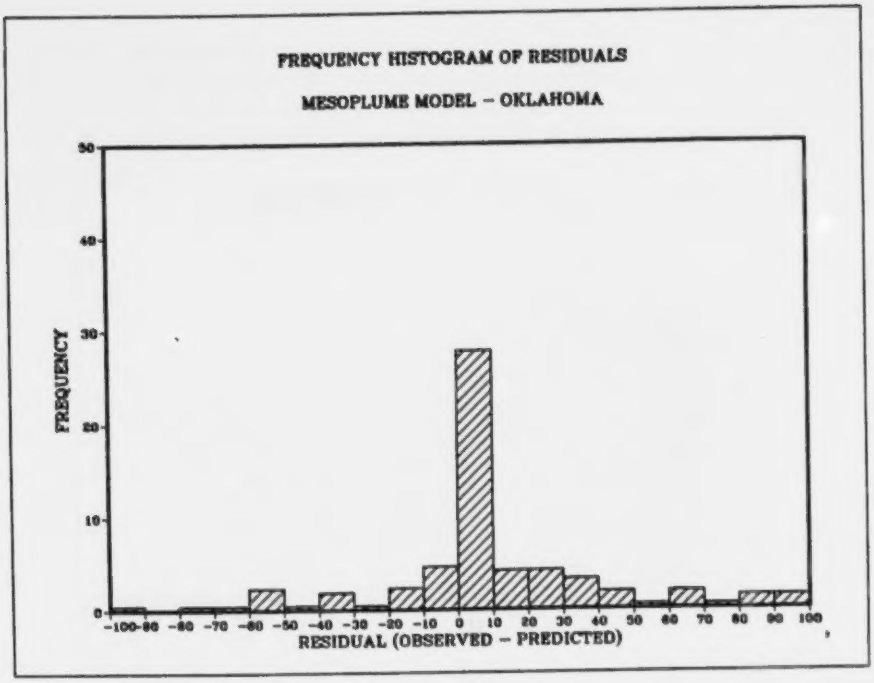


Figure 4-17. Frequency distribution of residuals at Oklahoma for MESOPLUME based on points paired in space and time ... residual range: -100 to 100 parts per 10^{15}.

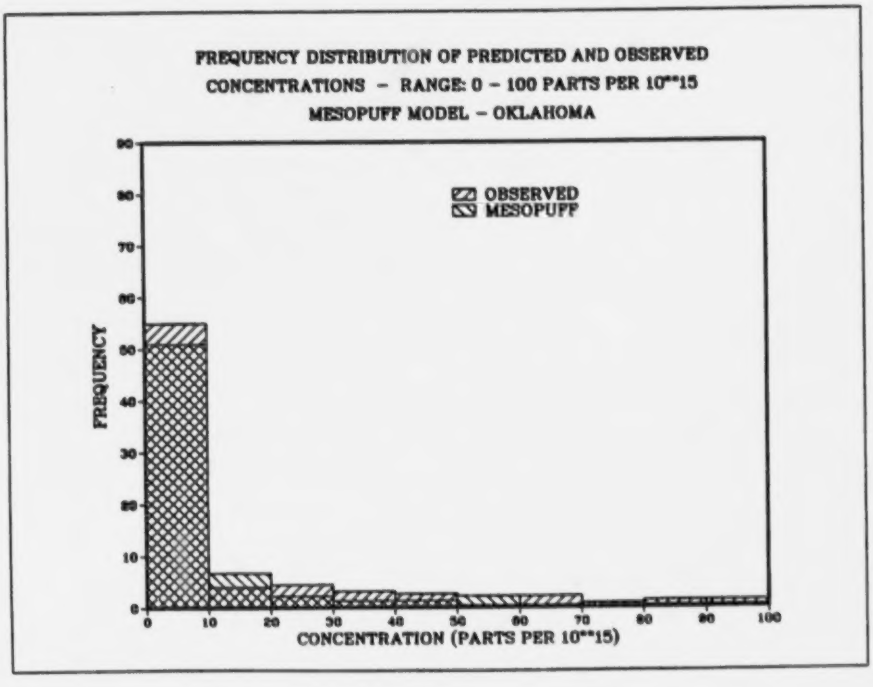


Figure 4-18. Frequency distribution of predicted concentrations at Oklahoma for MESOPUFF based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}.

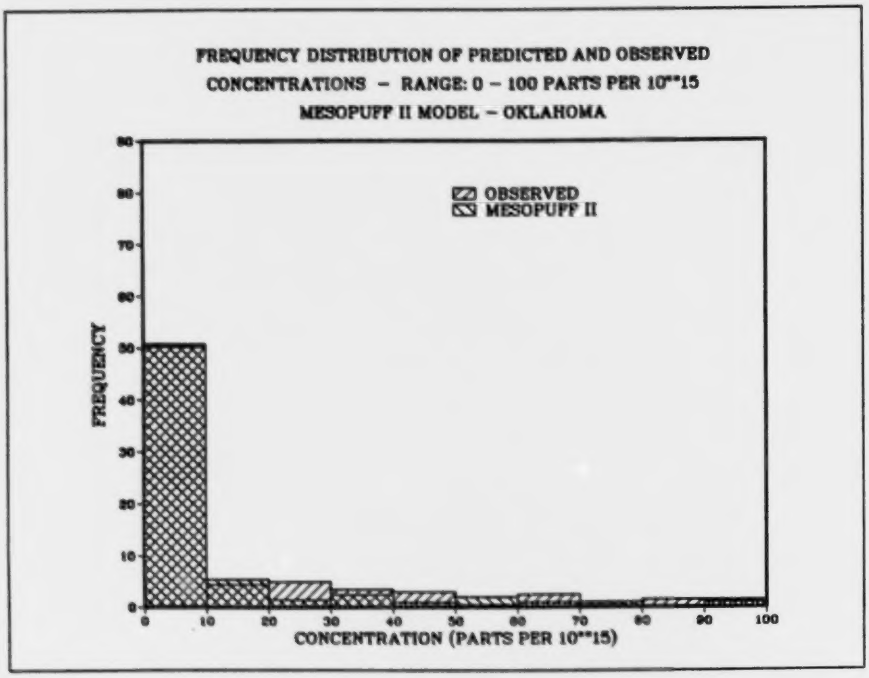


Figure 4-19. Frequency distribution of predicted and observed concentrations at Oklahoma for MESOPUFF II based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}.

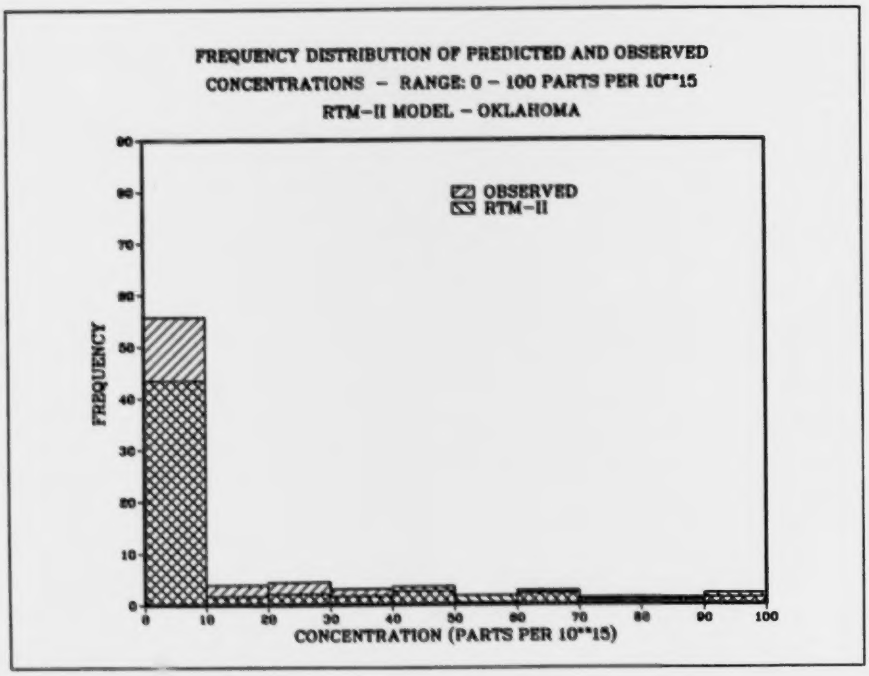


Figure 4-20. Frequency distribution of predicted and observed concentrations at Oklahoma for RTM-II based on points paired in space and time ... concentration range: 0 to 100 parts per 10^{15}.

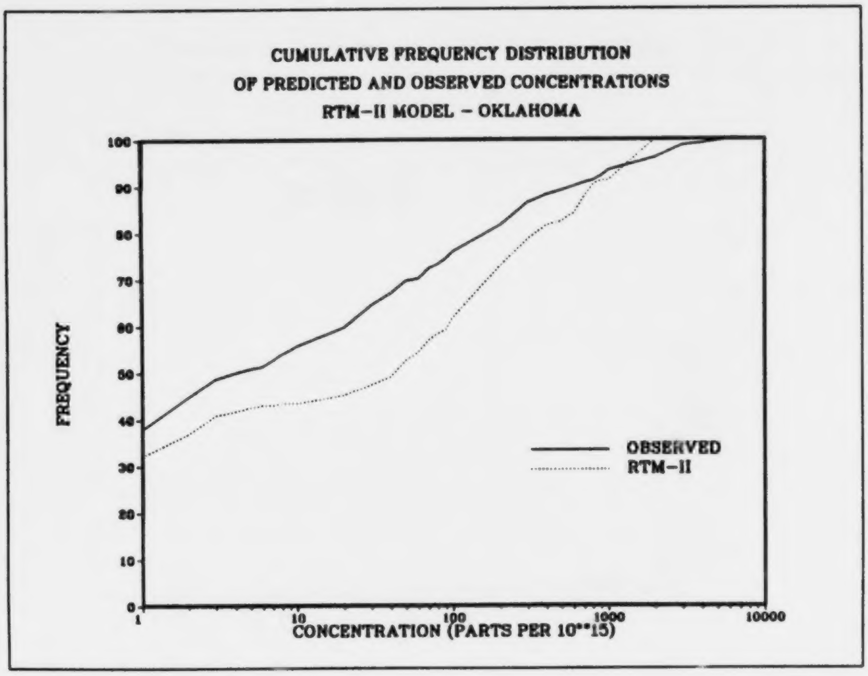


Figure 4-21. Cumulative frequency distributions of RTM-II predictions and observed concentrations at Oklahoma based on points paired in space and time.

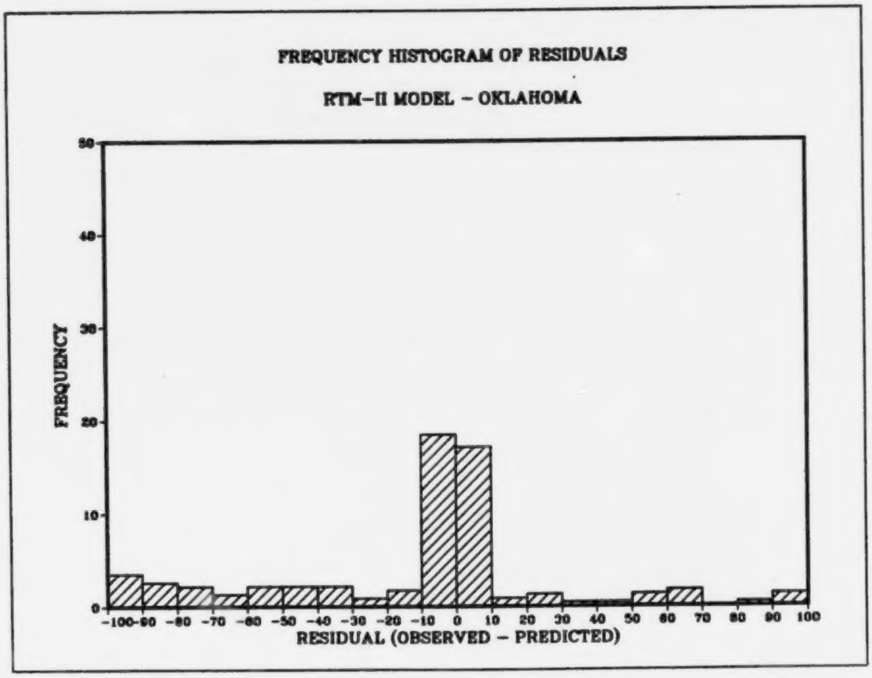


Figure 4-22. Frequency distribution of residuals at Oklahoma for RTM-II based on points paired in space and time ... residual range: -100 to 100 parts per 10^{15}.

Table 4-14. Summary of Results of Graphical Comparisons of Models with Oklahoma Data.

Model	Frequency Distribution of P and O (trend for 0-10 range)	Cumulative Frequency Plot (initial trend)	Frequency Histogram of Residuals (position of peak range)
MESOPUFF	O > P	P > O	0 - 10
MESOPLUME	P > O	P > O	0 - 10
MSPUFF	P > O	P > O	0 - 10
MESOPUFF II	P > O	P > O	0 - 10
ARRPA	P > O	P > O	0 - 10
RADM	P > O	P > O	0 - 10
RTM-II	O > P	O > P	-10 - 0

concentrations in parts per 10^{15} and "O" represents observed concentrations in the same units. Graphs like Figure 4-15 were examined to provide the first column, which reports whether the number of predicted ground concentrations exceeded the number of observed concentrations in the range of 0-10 parts/10^{15} (the "initial trend" of the frequency distribution). This entry determines whether a model predicts as many nonzero small concentrations as were observed. The cumulative frequency plots, like Figure 4-16 give the number of predicted and observed concentrations that were equal to or less than a given value on the horizontal axis. They indicate in what range of concentration values overprediction is occurring, and in what range the model tends to underpredict. The "initial trend" of this plot encompasses roughly the range of concentrations from 0 to 50 or 100 parts/10^{15}. The histogram of residuals (excess of observed over predicted without regard to magnitude of concentration) has a strong peak near zero for all models. Figure 4-17 is an example of this type of graph. Whether this peak occurs for small positive residuals or small negative residuals determines whether the model tends to underpredict ground concentrations regardless of the magnitude of the concentration (positive) or whether it tends to overpredict these concentrations (negative). The summary behavior in this table of all of the models except

RTM-II for the first two types of graphical comparisons indicates overprediction of plume ground concentrations for values in the range of 0-50 parts/10^{15}. However, the positive values for the peak of the frequency histogram of residuals for these models indicates overall underprediction of values when larger observed concentrations are included. RTM-II clearly exhibits the opposite behavior. It predicts too few small concentrations, but too many large ones. The first column for MESOPUFF indicates that this model tends to underpredict plume ground concentrations for the smallest range of observed values, but the second column entry for MESOPUFF indicates overprediction of concentrations in the range of 10-50.

The predictions of MESOPLUME are presented first in Figures 4-15 to 4-17 to show the typical prediction of the models. The other graphs present the sole deviations from the trends represented by the MESOPLUME predictions. (Appendix E contains the full set of graphs for each model.) Note that the RTM-II graphs show opposite trends to the "standard" MESOPLUME graphs.

4.3.2.2. Savannah River Plant Data Base

Table 4-15 has been developed for the SRP graphs to parallel Table 4-14 for the Oklahoma graphs. The "standard behavior" graphs for the SRP data cases are Figures 4-23 to 4-25, and graphs of deviations from the standard are presented in Figures 4-26 to 4-30.

Three of the models (MESOPLUME, MSPUFF and RADM) exhibit the same behavior in all three types of graphical comparisons as they did for the Oklahoma data. The MTDDIS model underpredicts plume ground concentrations as shown by the first two types of graphs, but overpredicts large observed values. The mixed results summarized in Table 4-15 for MESOPUFF II and RTM-II show that these two models have predicted and observed ground concentration distributions that are more nearly in agreement. RTM-II predicts the same number of concentrations as observed in the 0-10 range, but the cumulative number of predictions above 10 parts/10^{15} exceeds the observed number. However, for RTM-II, the peak of the residuals is above zero indicating some underprediction of large concentrations. MESOPUFF II predicts fewer small concentrations than were observed, but its cumulative frequency distribution also exceeds the observed. Its tendency toward some underprediction of large concentrations is shown by the occurrence of a positive value for the peak of

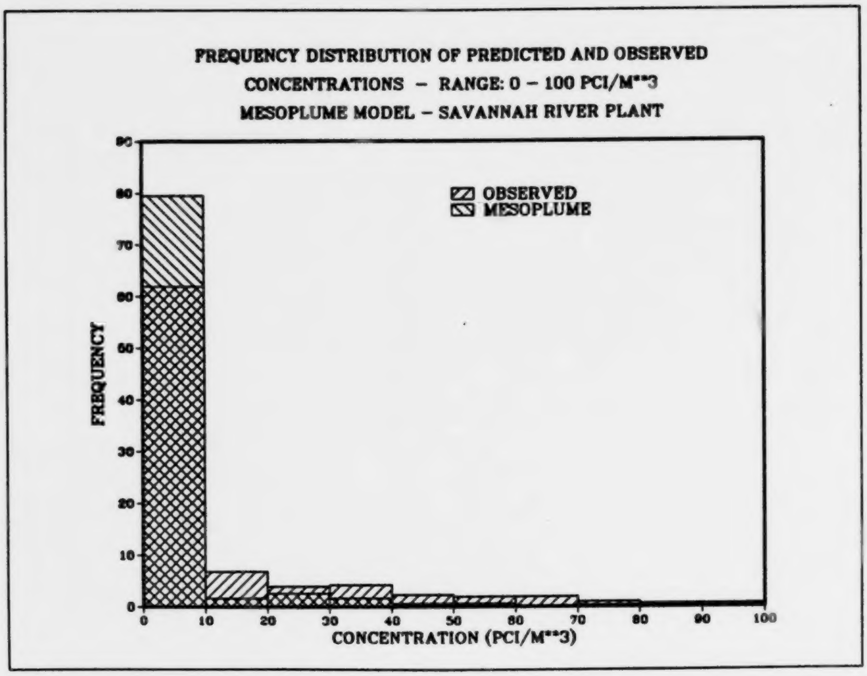


Figure 4-23. Frequency distribution of predicted and observed concentrations at Savannah River Plant for MESOPLUME based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3.

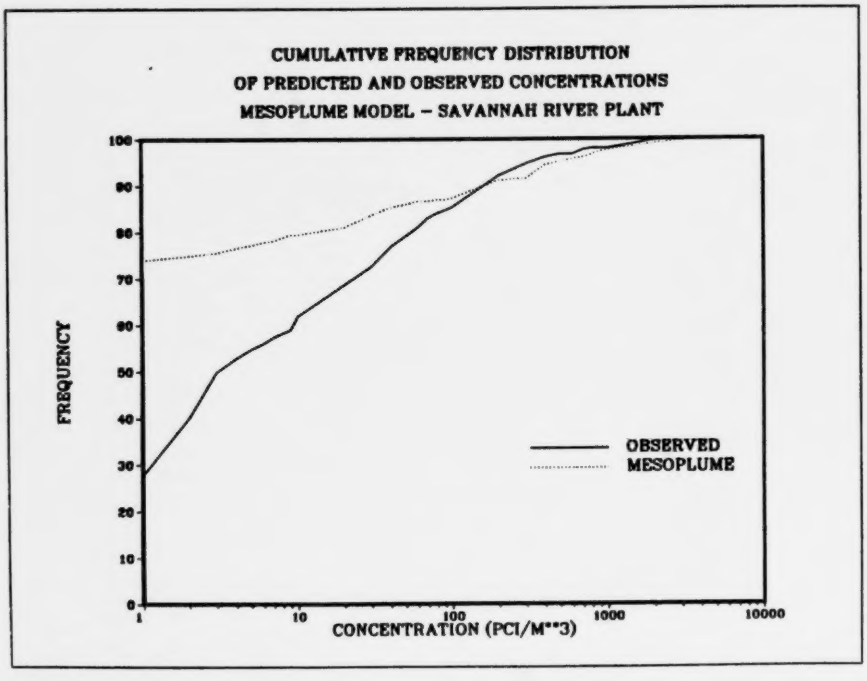


Figure 4-24. Cumulative frequency distributions of MESOPLUME predictions and observed concentrations at Savannah River Plant based on points paired in space and time.

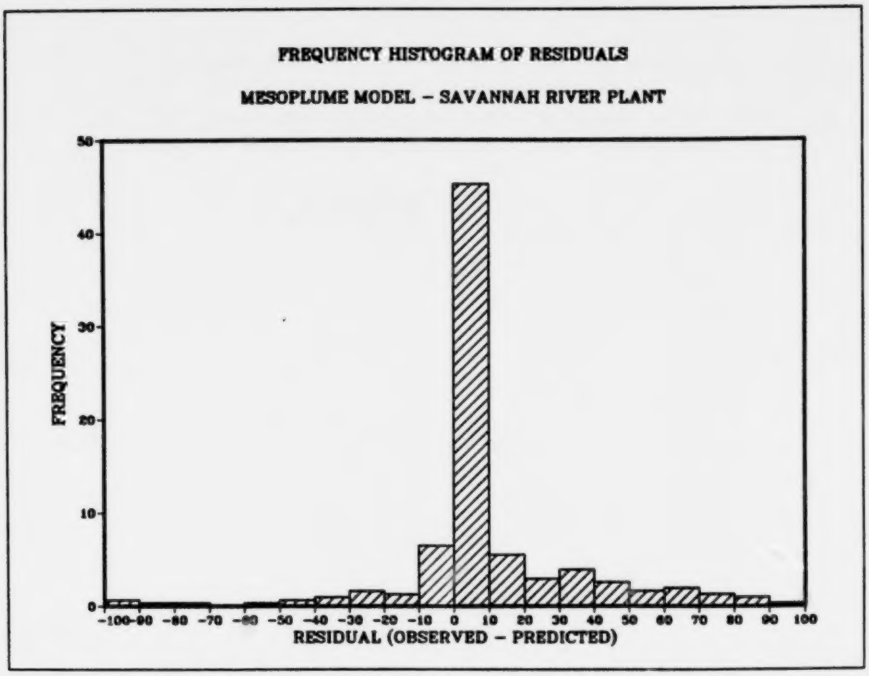


Figure 4-25. Frequency distribution of residuals at Savannah River Plant for MESOPLUME based on points paired in space and time ... residual range: -100 to 100 pCi/m^3.

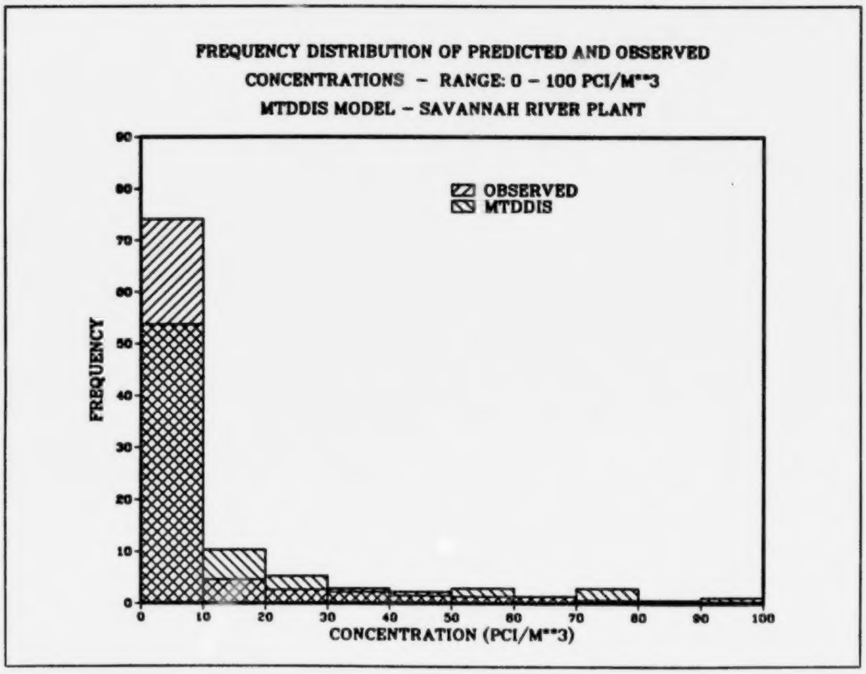


Figure 4-26. Frequency distribution of predicted and observed concentrations at Savannah River Plant for MTDDIS based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3.

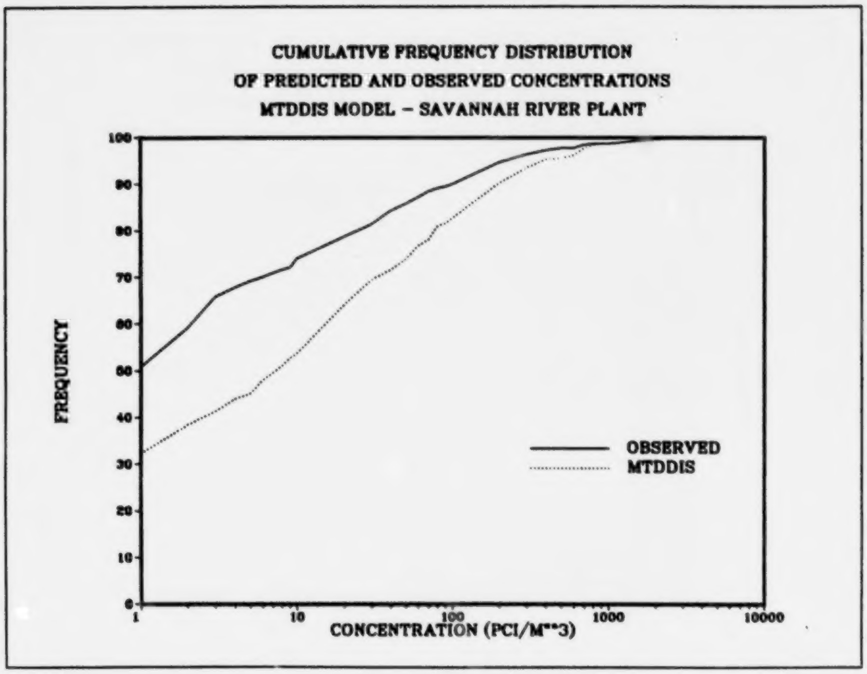


Figure 4-27. Cumulative frequency distribution of MTDDIS predictions and observed concentrations at Savannah River Plant based on points paired in space and time.

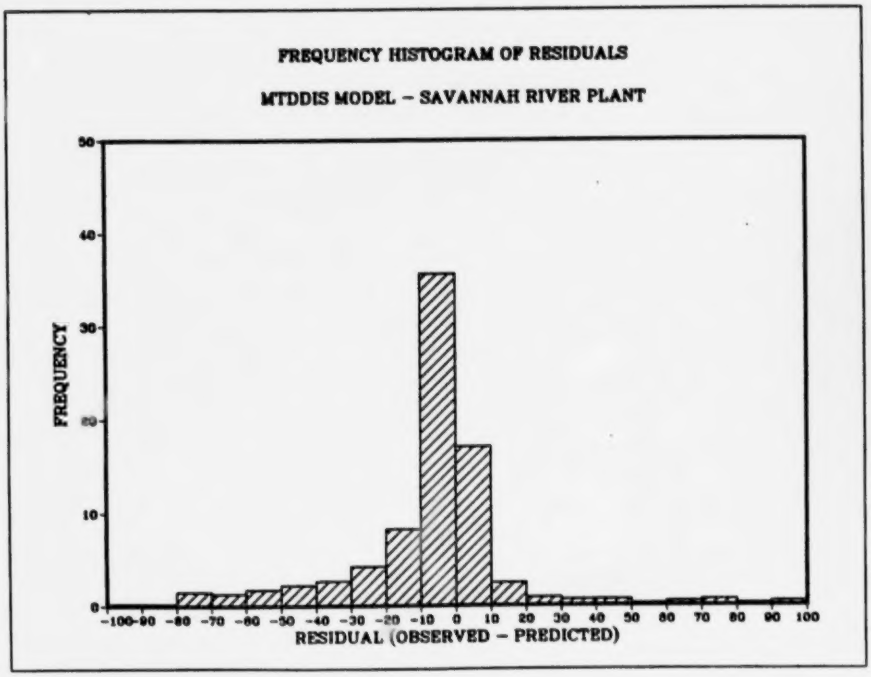


Figure 4-28. Frequency distribution of residuals at Savannah River Plant for MTDDIS based on points paired in space and time ... residual range: -100 to 100 pCi/m^3.

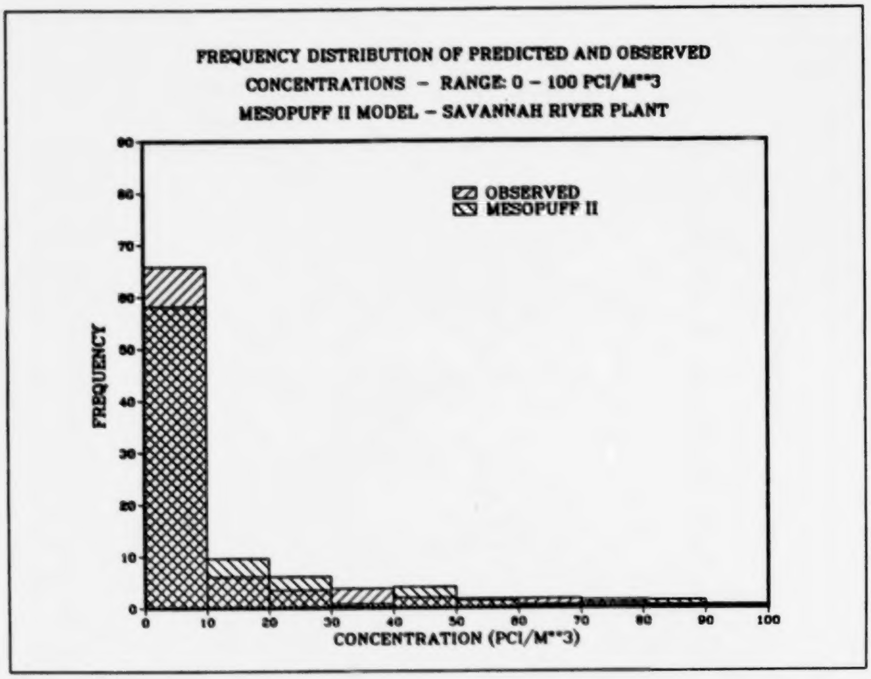


Figure 4-29. Frequency distribution of predicted and observed concentrations at Savannah River Plant for MESOPUFF II based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3.

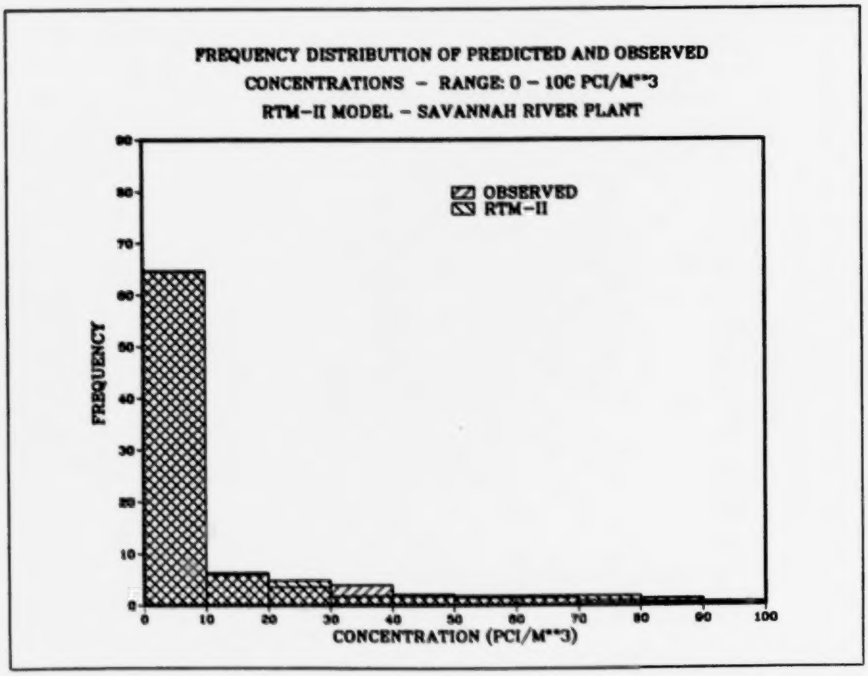


Figure 4-30. Frequency distribution of predicted and observed concentrations at Savanna River Plant for RTM-II based on points paired in space and time ... concentration range: 0 to 100 pCi/m^3.

Table 4-15. Summary of Results of Graphical Comparisons of Models with Savannah River Plant Data.

Model	Frequency Distribution of P and O (initial trend, 0-10 range)	Cumulative Frequency Plot (initial trend)	Frequency Histogram of Residuals (position of peak range)
MESOPUFF	P > O	P > O	0 - 10
MESOPLUME	P > O	P > O	0 - 10
MSPUFF	P > O	P > O	0 - 10
MESOPUFF II	O > P	P > O	0 - 10
MTDDIS	O > P	O > P	-10 - 0
RADM	P > O	P > O	0 - 10
RTM-II	P = O	P > O	0 - 10

the residuals. MESOPUFF predicted fewer concentrations in the 0-10 parts/10^{15} range than were observed for the Oklahoma data; however, the opposite trend was found when comparisons were made with the SRP data.

Again, the MESOPLUME graphs that are presented in Figures 4-23 to 4-25 represent the common behavior of the models. The remaining graphs present deviations from that common behavior. As may be seen, the MTDDIS model provides the opposite behavior to the "standard" MESOPLUME prediction characteristics. MTDDIS reveals a much greater plume spreading than shown by the other six models and also has a tendency to overpredict with respect to the ground-level concentration data as well. No other noteworthy deviation can be systematically identified from Table 4-15.

SECTION 5

DIAGNOSTIC EVALUATION BASED ON MODEL/DATA COMPARISONS

5.1. INTRODUCTION

In this section, additional graphical and tabular comparisons of model predictions with observed data will be presented. The specific features of the predictive behavior of the models will be traced to model assumptions and experimental constants. Although the AMS statistics provide an objective method for evaluating the predictive accuracy of the models, they do not provide information that can be used to identify the causes of model/data discrepancies. For example, the predicted plume from a model might be approximately the right shape and might contain realistic concentration values, but its ground concentration pattern might simply be shifted from the observed plume. Such a directional inaccuracy of the predicted plume could arise from a combination of errors in the predicted wind field from the meteorological preprocessor and from the coarse spatial resolution of the input meteorological data. In such a case the AMS statistics will only show model/data disagreement, but cannot indicate that the pattern is correct, yet in an incorrect location. A further analysis of the model predictions has been carried out to help provide more insight into the causes of the model/data discrepancies.

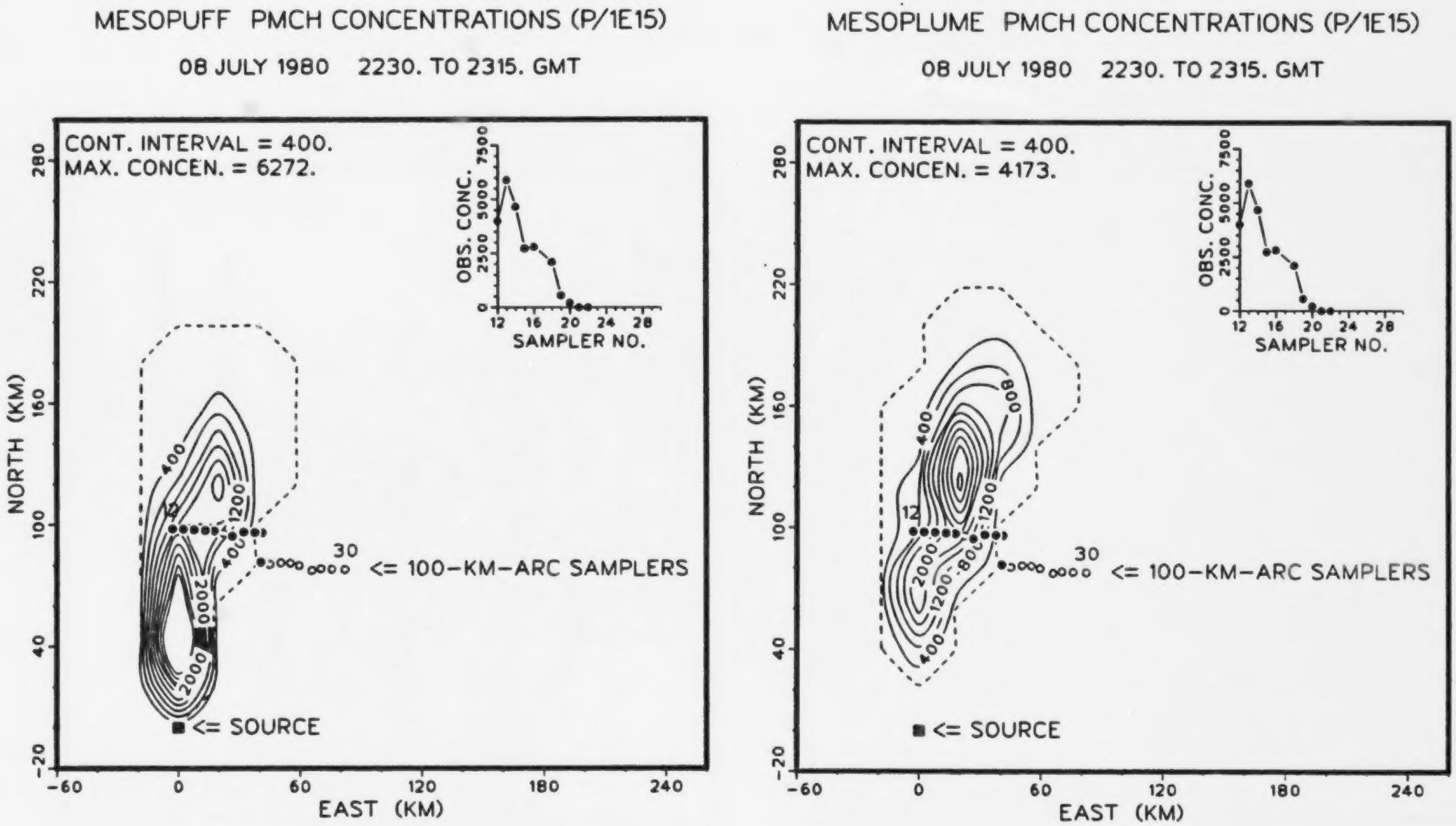
Section 5.2 presents a subset of the plume concentration isopleth plots that have been prepared for each of the 21 sampling periods in the Oklahoma data base. Section 5.3 also presents a representative sample of plume concentration graphs prepared for the 65 sampling periods in the Savannah River Plant data base. These graphs provide a clear comparison of the location and width of the predicted plumes relative to the data. They allow several general conclusions to be made concerning the relationship between model predictive performance and modeling assumptions.

In Section 5.4, a quantitative pattern comparison method is applied to the Oklahoma data. This method was not applied to the Savannah River Plant data cases since the samplers are too widely spaced to allow the same degree of quantitative analysis. Tables have been prepared comparing predicted and observed values of (a) plume width, (b) azimuth angle of transport through sampler arcs and (c) time of arrival at the arcs. These comparisons provide new insight into the AMS statistical results for each model. In that section, the accuracy of several common modeling assumptions is discussed on the basis of the pattern comparison results.

The final section, Section 5.5, presents a simplified yet quantitative pattern comparison method for the Savannah River Plant data, consistent with the wide spacing of samplers in that experiment. Tables of maximum predicted concentrations for each averaging period for both data bases are then presented. These tables allow the models to be intercompared on the basis of whether they tend to predict ground concentration patterns that are always higher or lower than those of the other models. Finally, the theoretical assumptions for each of the eight models are examined in light of the predictive performance of the model. Section 5.5 ends with a discussion synthesizing all the diagnostic comparison methods described in this Section 5. For each model, the specific assumptions responsible for the model's predictive performance relative to observations are discussed, and causes of inaccuracies are identified where possible.

5.2. COMPARISON OF PREDICTED CONCENTRATION ISOPLETHS WITH DATA AT OKLAHOMA

Figures 5-1 through 5-12 illustrate the type of isopleth plot prepared for each model for each of the 21 sampling periods in the Oklahoma data base. The remainder of these graphs appear as Figures E-12 through E-95. The model name, date and time-averaging period are printed at the top of the graph. The 20 km grid steps are indicated by tick marks on the axes, and the source is located at the origin, (0,0). The contour interval used for the isopleth map is printed in the upper left corner in parts/10^{15}, and isopleth values are presented in those units in the figure. The maximum predicted concentration found on the prediction grid of size 34 x 34 (20 km by 20 km grid spacing) is listed beneath the designation of contour interval. (Since a constant contour interval is used, and no more than 10 contours were permitted on a graph, the peak concentration is often much larger than that of the innermost contour shown.) The black dots refer to samplers that were turned on during the



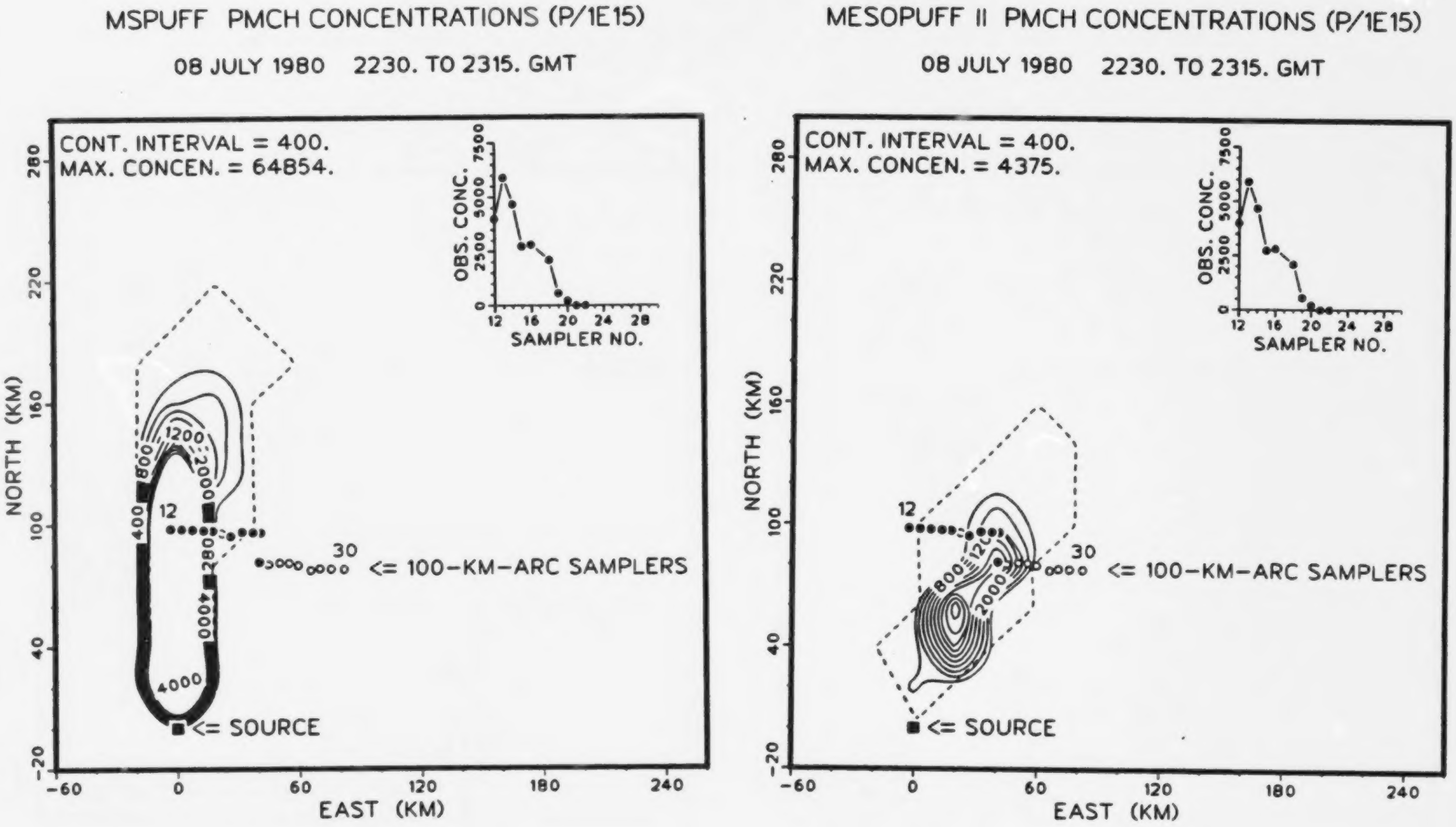
MESOPUFF PMCH CONCENTRATIONS (P/1E15)

08 JULY 1980 2230. TO 2315. GMT

MESOPLUME PMCH CONCENTRATIONS (P/1E15)

08 JULY 1980 2230. TO 2315. GMT

Figure 5-1. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions.



MSPUFF PMCH CONCENTRATIONS (P/1E15)

08 JULY 1980 2230. TO 2315. GMT

MESOPUFF II PMCH CONCENTRATIONS (P/1E15)

08 JULY 1980 2230. TO 2315. GMT

Figure 5-2. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) MSPUFF predictions, (right) MESOPUFF II predictions.

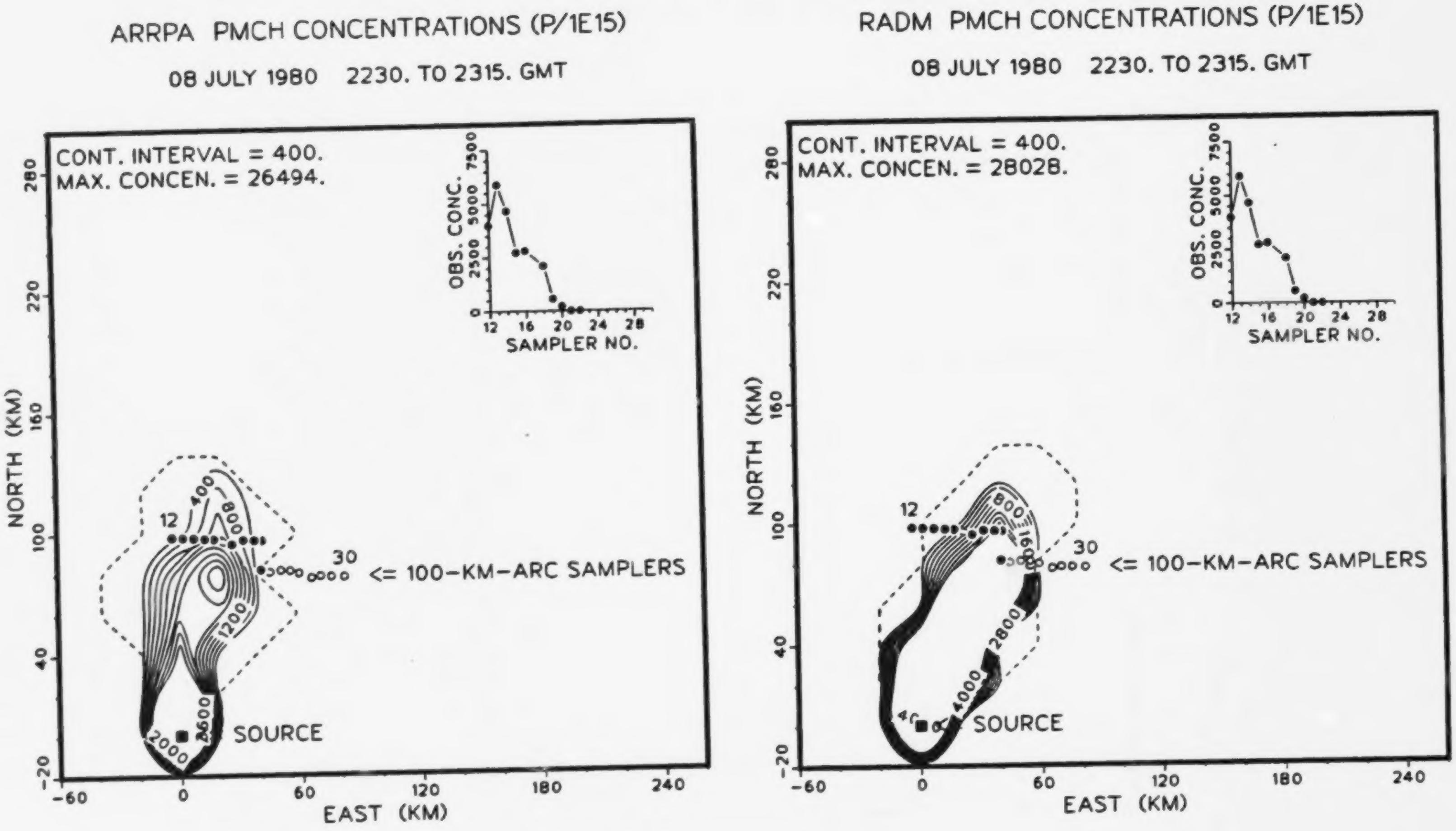


Figure 5-3. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... (left) ARRPA predictions, (right) RADM predictions.

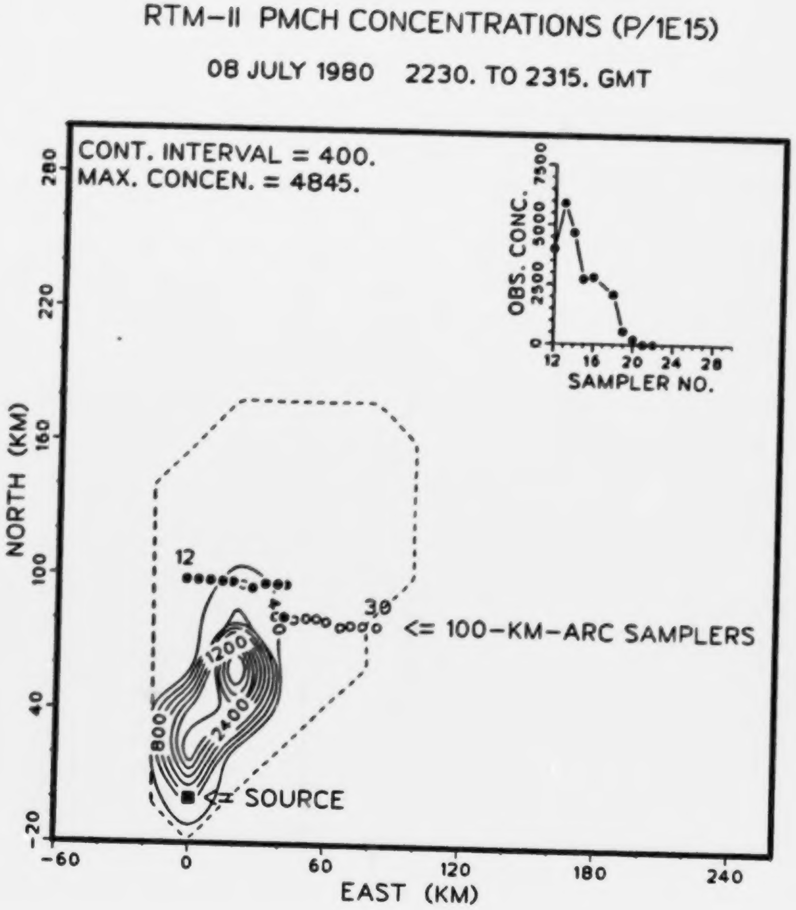
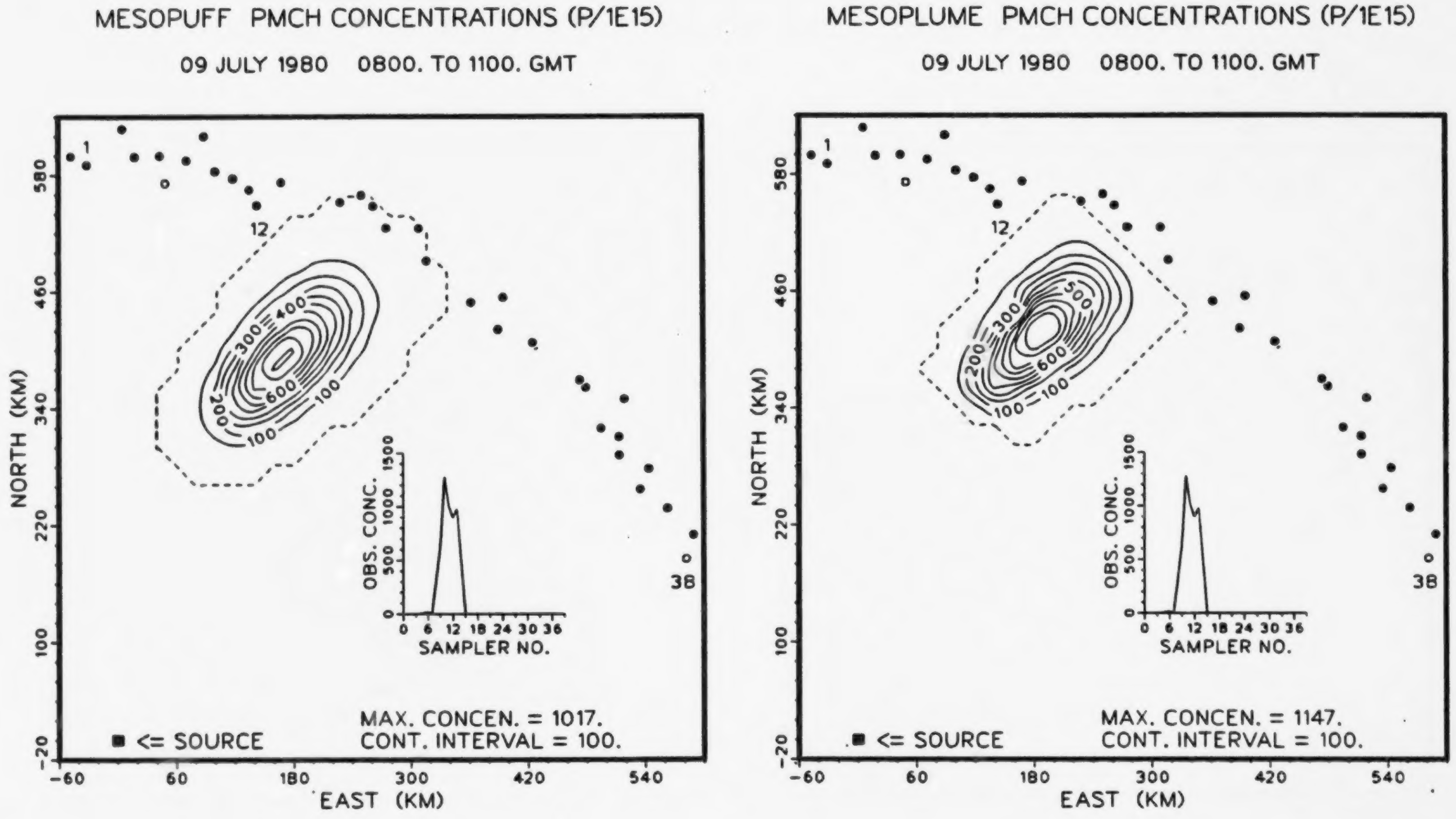


Figure 5-4. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 8, 1980 (2230 to 2315 GMT) ... RTM-II predictions.



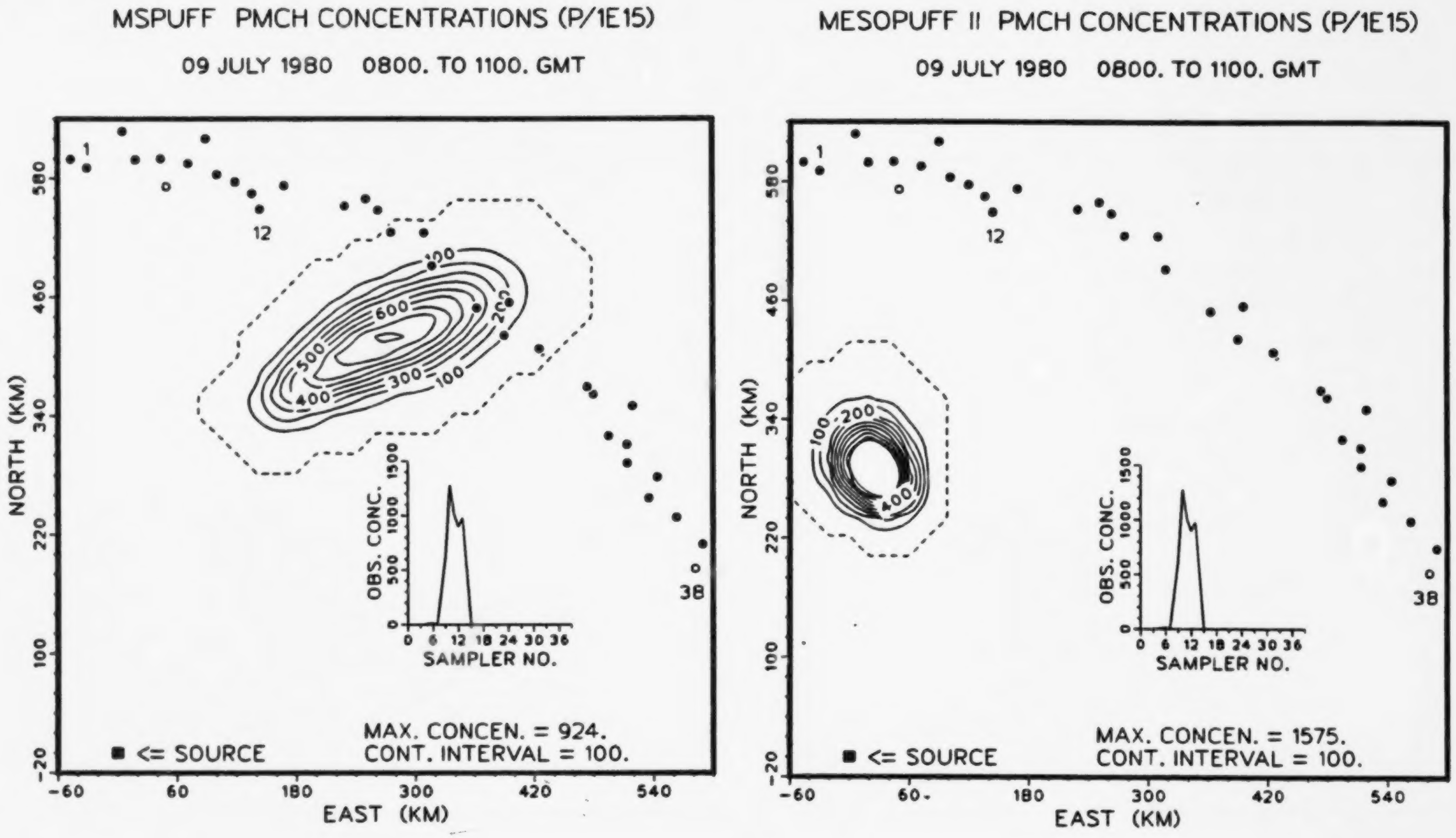
MESOPUFF PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

MESOPLUME PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

Figure 5-5. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions.



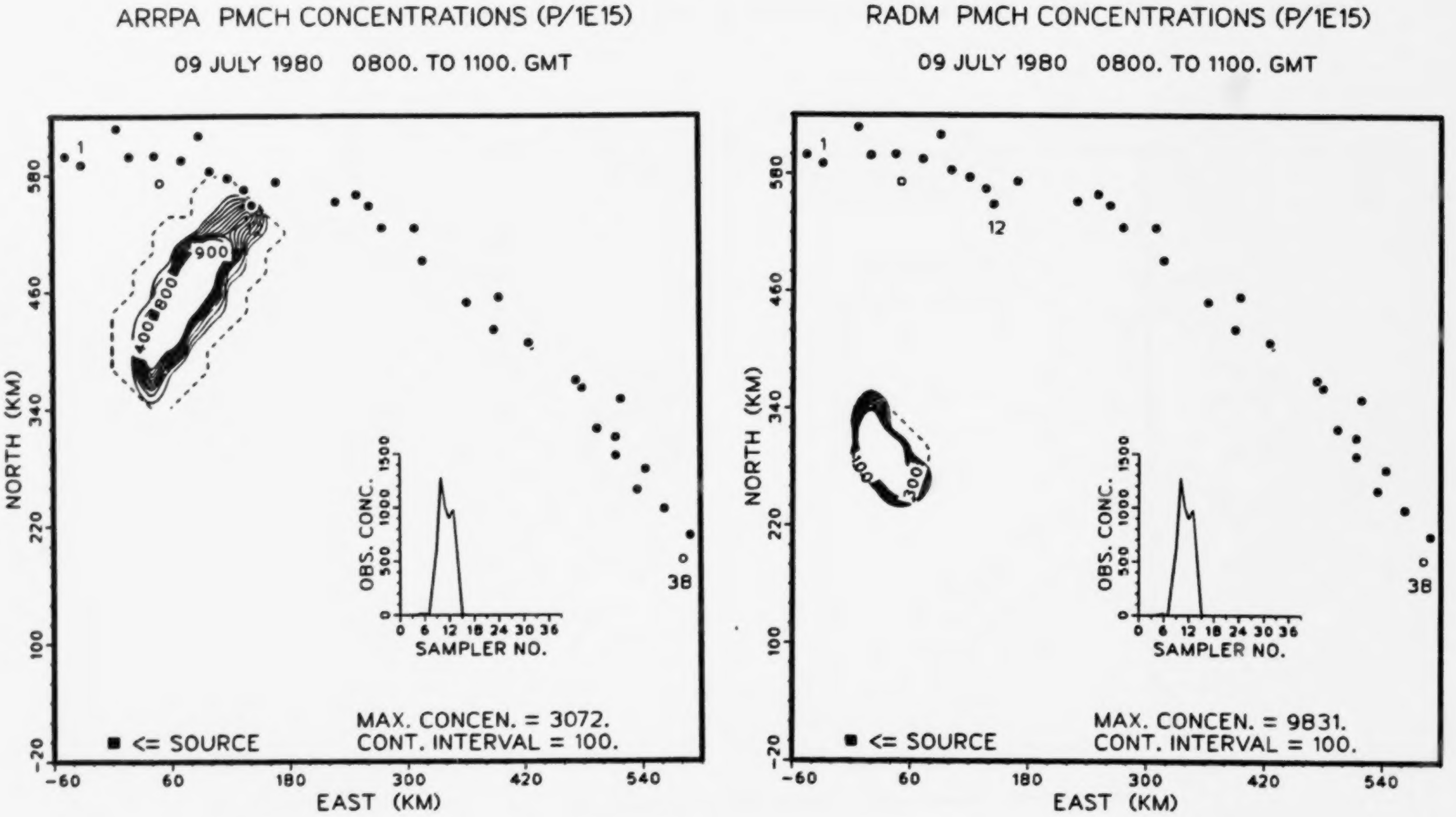
MSPUFF PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

MESOPUFF II PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

Figure 5-6. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ...
(left) MSPUFF predictions, (right) MESOPUFF II predictions.



ARRPA PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

RADM PMCH CONCENTRATIONS (P/1E15)

09 JULY 1980 0800. TO 1100. GMT

Figure 5-7. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... (left) ARRPA predictions, (right) RADM predictions.

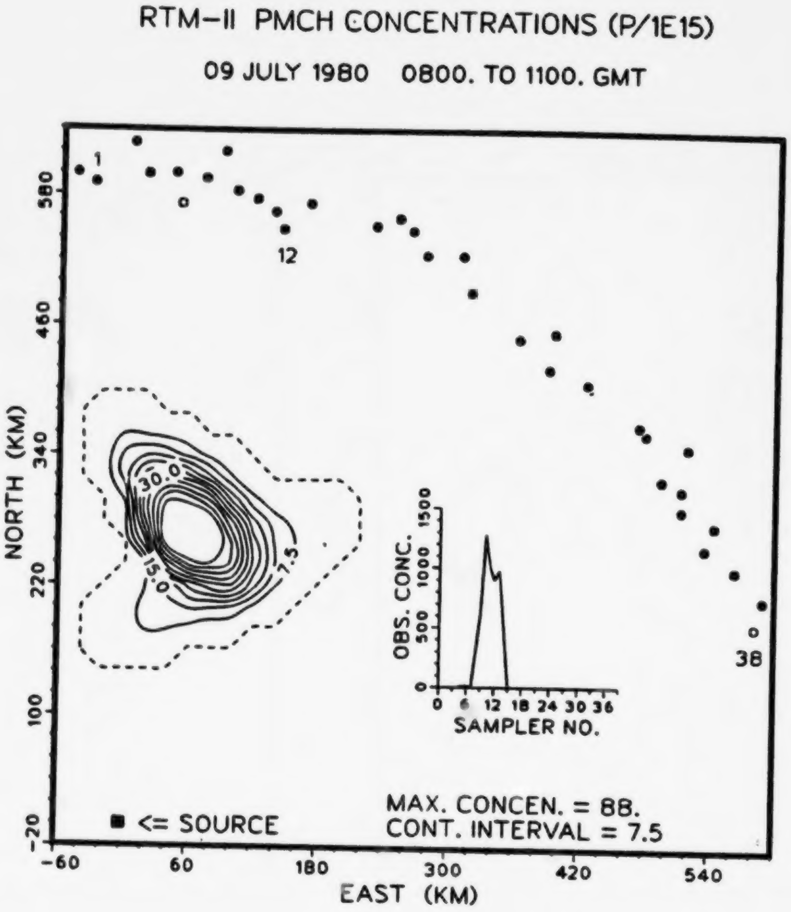


Figure 5-8. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 9, 1980 (0800 to 1100 GMT) ... RTM-II predictions.

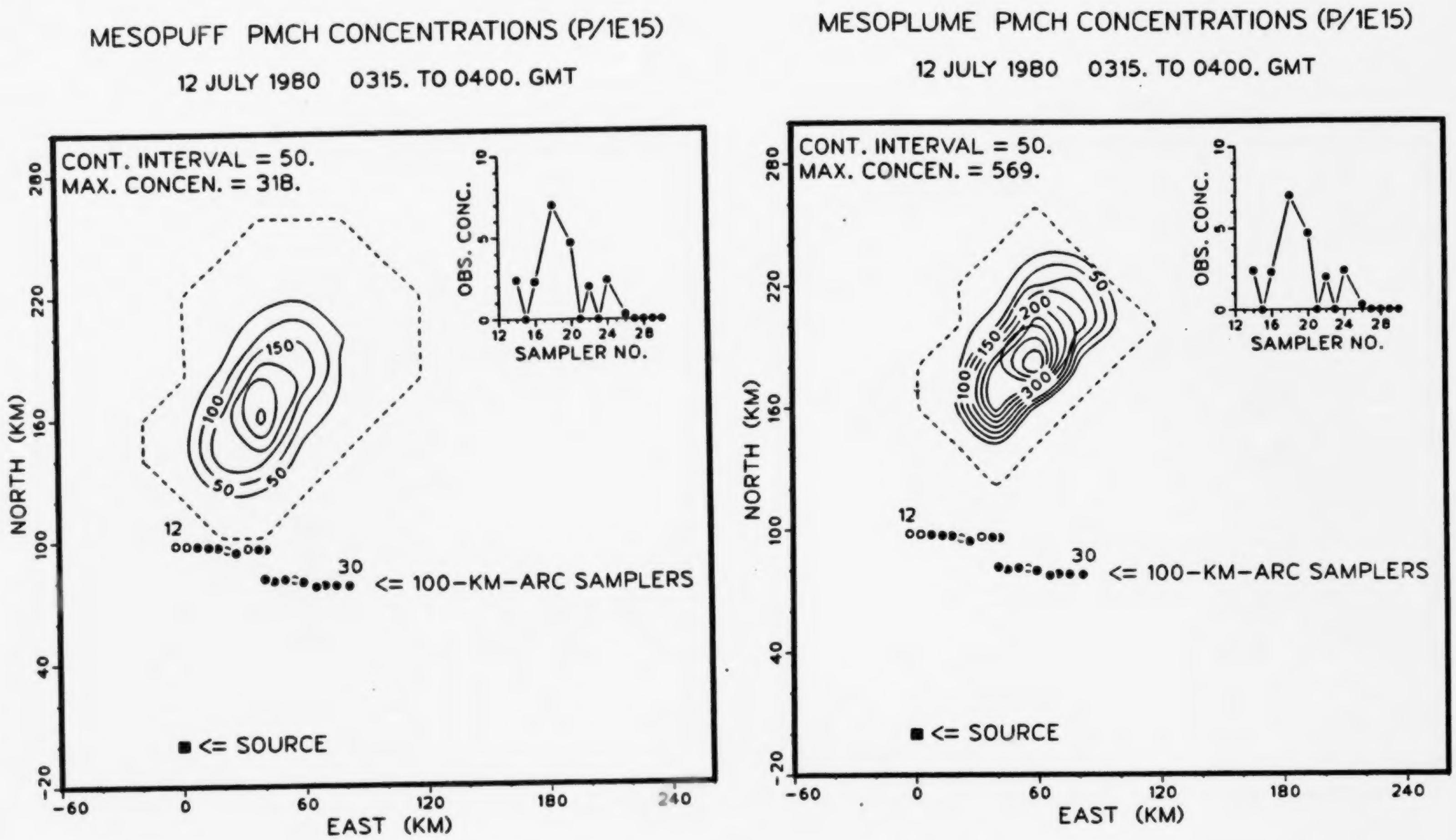


Figure 5-9. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) MESOPUFF predictions, (right) MESOPLUME predictions.

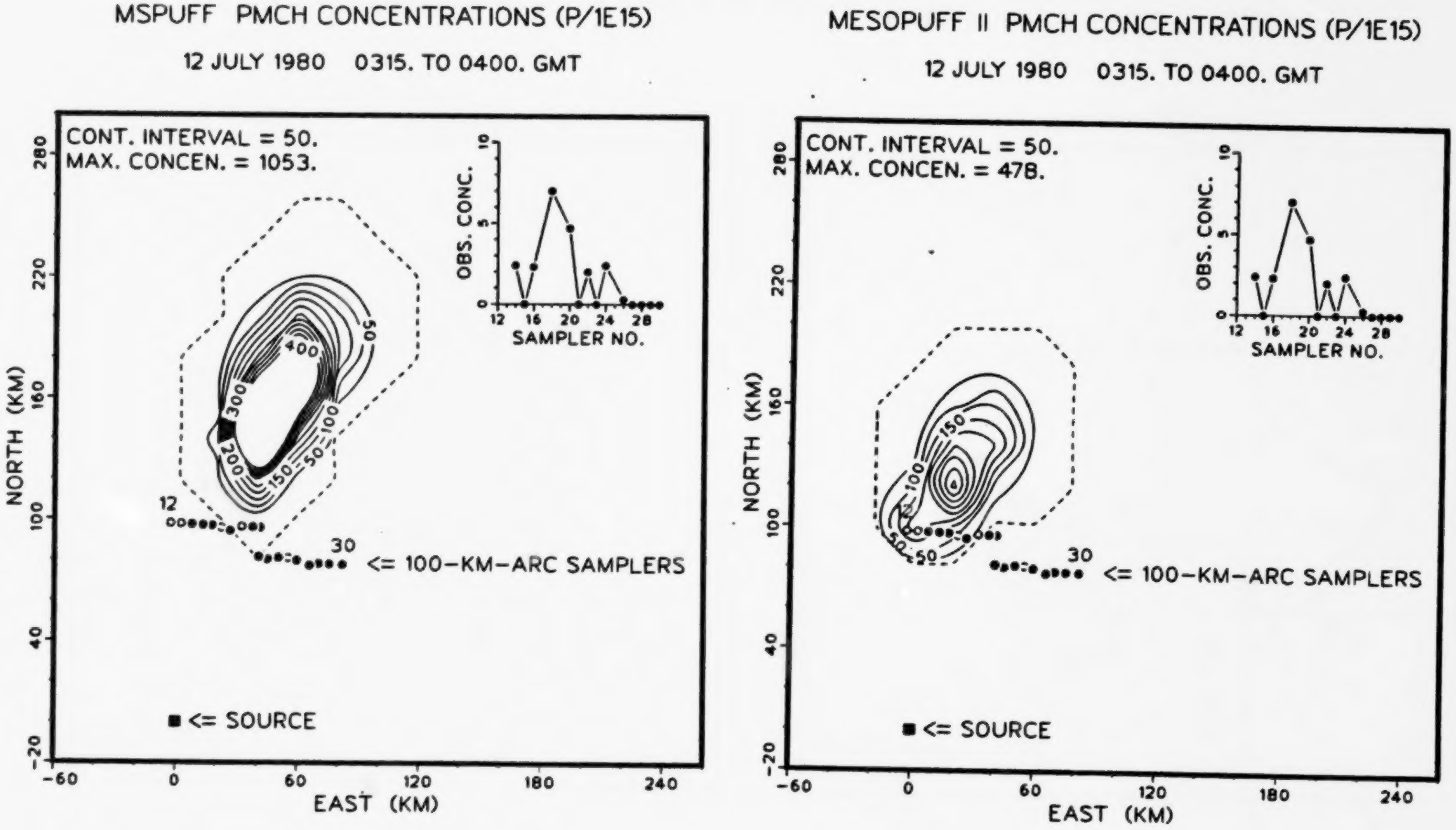


Figure 5-10. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) MSPUFF predictions, (right) MESOPUFF II predictions.

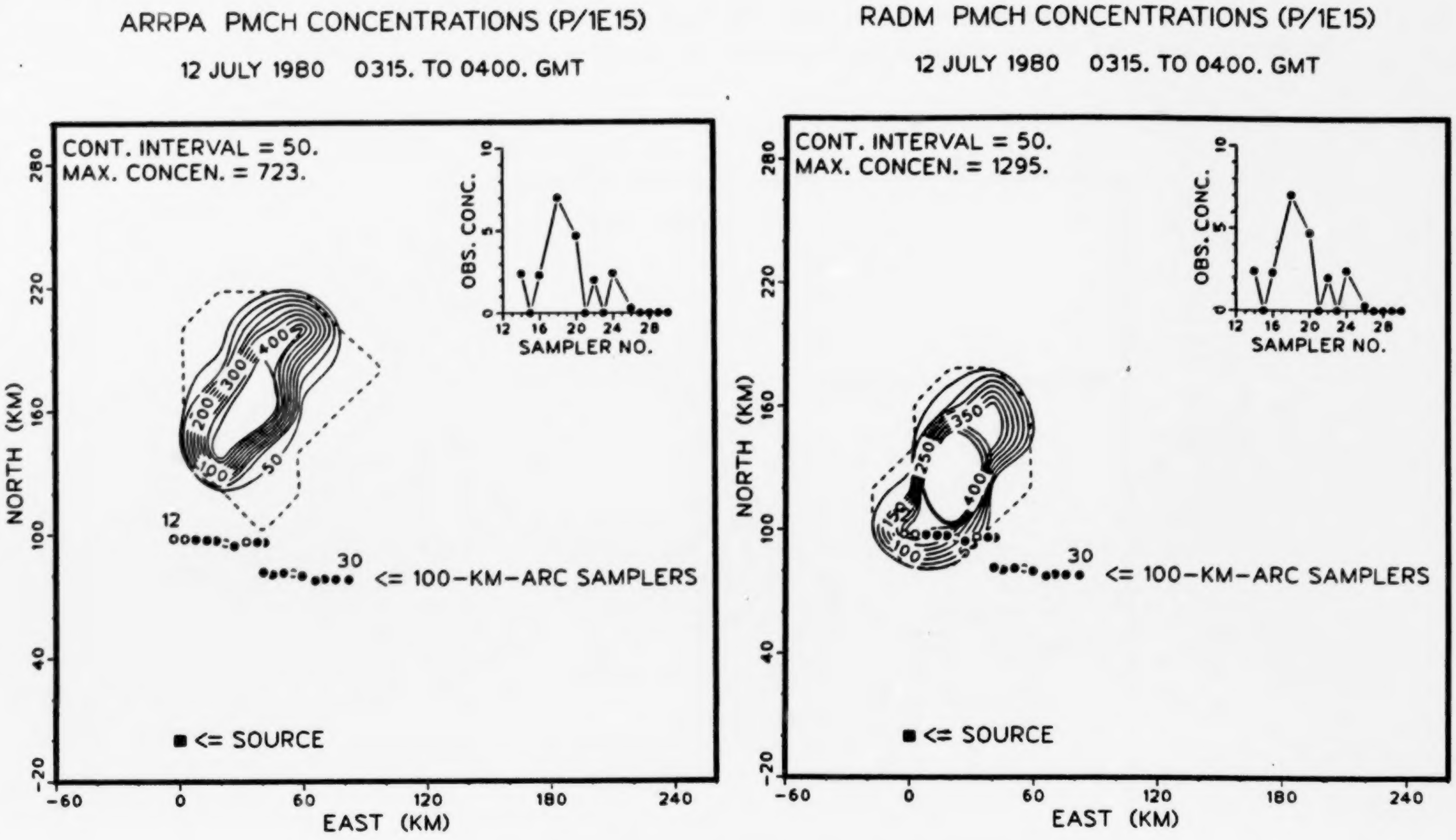
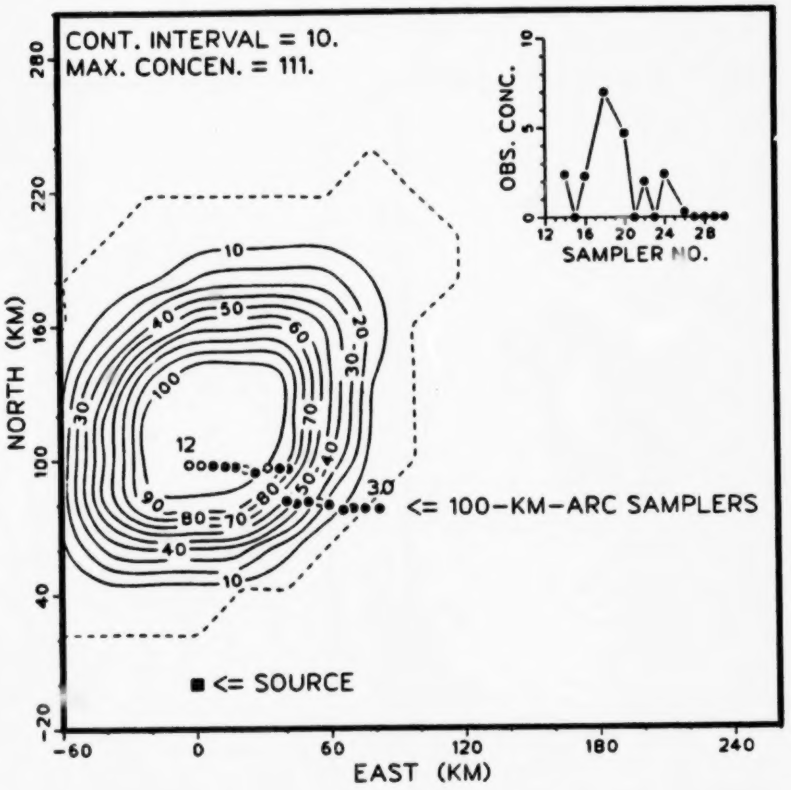


Figure 5-11. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... (left) ARRPA predictions, (right) RADM predictions.



RTM-II PMCH CONCENTRATIONS (P/1E15)

12 JULY 1980 0315. TO 0400. GMT

Figure 5-12. Isopleth plot of ground-level concentrations for the Oklahoma experiment of July 12, 1980 (0315 to 0400 GMT) ... RTM-II predictions.

measurement period cited in the figure title; the unfilled dots refer to sampling stations that were not in operation during that period. Only a few samplers are labeled to avoid obscuring important contour features. The identification number of intermediate samplers is determined by counting from the labeled ones. The dotted line in the isopleth plots refers to the smallest curve that encloses non-zero predictions made by the model. (It passes through the points with zero predicted concentration nearest the region with non-zero values.) Finally, in the upper right corner is drawn a graph of the observed concentration values at each active sampler across the 100 or 600 km arc, joined by straight lines. The x-axis is the sampler identification number. (Equal spacing on this axis does not indicate equal distances between samplers.) The y-axis units are parts/10^{15}.

The interpolation scheme used to prepare the isopleth plots assumes, for any point inside a grid cell, a linear variation of concentration based on predictions made at the four nearest grid nodes. This method produces concentration values within a cell that are always smaller than the largest value at the four neighboring grid points. It should be recognized, however, that model predictions with a finer grid spacing would likely reveal a larger value within a cell.

Figures 5-1 to 5-4 present isopleth plots of the seven model predictions for the time period representing the peak observed concentration at the 100 km arc for the July 8 experiment. This time period represents the 45-minute averaging period 2230 to 2315 GMT; the perfluorocarbon release took place from 1900 to 2200 GMT. The actual transport time from the source to the 100 km arc was about five hours. Figures 5-5 to 5-8 present similar isopleths for the 600 km arc during the period expected to be the time of peak observed concentration, 0800 to 1100 GMT on July 9, 1980. (As the data indicated, the concentration peak arrived earlier before the samplers were turned on. This may have been due to rapid transport by an elevated nocturnal jet.) On July 11, 1980 a second three-hour release was made from 1900 to 2200 GMT. Figures 5-9 to 5-12 present predictions for the concentration averaging period 0315 to 0400 GMT on July 12, 1980. Peak observed concentrations on the 100 km arc did not occur during this time period. That period represents the last one for which measurements at the 100 km arc were taken. Isopleth plots for that time period were chosen for presentation here because it best illustrates the differences in the concentration patterns predicted by the models.

Two special aspects of the above graphs should be explained. In Figures like 5-3 and 5-4, some of the isopleths pass upwind of the source (with the wind from the southwest). This is an non-physical feature of the isopleth program resulting from linear interpolation between the source grid point and the zero concentration values at the nearest upwind grid points. Such nonzero upwind values should be ignored as an artificial result of the interpolation method used. Also, in Figure 5-9, the 350 parts/10^{15} contour crosses the 400 parts/10^{15} contour near the peak of the concentration pattern. This is also obviously not a physical effect, but results from the isopleth program logic that develops each isopleth without checking nearby ones to eliminate crossings or to make other adjustments. Although more sophisticated isopleth-generating programs are available, the method used provides adequate interpretive information for this study.

An examination of the above figures along with those in Appendix E reveals the following conclusions:

(a) All sampler values representing maximum predicted values are offset in time from the observed maximum values. For the July 8 experiment, a 1-3 hour time offset is typical at the 100 km arc and a 2-8 hour time offset is representative of transport to the 600 km arc. Some models predict plumes that transport slower and others faster than the observed data at the 100 km arc for the July 8 experiment. All models predict a slower transport to the 600 km arc as compared with the observed plume. The time lag/lead problem also exists for the 100 km arc in the July 11 experiment. The MESOPUFF and MESOPLUME model predictions lead the observed plume. Both models employ MESOPAC as the meteorological preprocessor. For the other models, the time lag is about one hour to the 100 km arc. Isopleths for the peak observation period for the second Oklahoma experiment (not presented here) may be found in Appendix E.

(b) All model predictions are offset from the data in space. Differences are largely dependent on the choice of meteorological preprocessor. Models with the same preprocessor (MESOPUFF and MESOPLUME using MESOPAC; and MESOPUFF II, RADM, and RTM-II using MESOPAC II) have the same general orientation of their isopleth pattern both in space and time. The horizontal and vertical dispersion algorithms used by each plume model largely create the differences in patterns within the same wind field.

(c) Comparisons of the predicted peak concentrations and horizontal spreading among the models reveal large differences. For these two Oklahoma cases, the RTM-II model tends to predict the greatest horizontal spreading with the RADM model predicting the least spreading. In terms of peak concentrations, the RADM model appears to predict the largest ground concentrations and the RTM-II model predicts the lowest peak values. Conservation of mass principles lead us to expect that models that predict the highest ground levels would have the least spreading. Clearly, a third dimension (the vertical) must play a role, yet, the vertical dimension may have a lesser effect since full mixing up to the mixed layer probably is occurring in each model at these 100 km and 600 km distances.

Some of this systematic behavior can be supported by means of a table of predicted and observed concentrations. Table 5-1 presents a listing of the predictions of the models and the observed ground-level data for both Oklahoma experiments as functions of ranges of values in parts/10^{15}. All time periods, both arcs, and both experiments are considered together here. All predicted/observed pairs (based on an original pairing in space and time) were employed in the preparation of this table. (All pairs with both members equal to zero were eliminated, however).

The large spatial and temporal offset of the patterns leads to large numbers of predicted or observed values that are zero in this table. If values greater than 300 parts/10^{15} are considered, the RADM model has many more of these large values than the other models, supporting the conclusion that RADM tends to overpredict the high concentrations. In fact, examination of Table 5-1 reveals that 42 observed values are greater than 200 parts/10^{15}. However, all models predict more than 51 values greater than 200 parts/10^{15}, with RADM predicting 68. Note also that RTM-II has the fewest number of predicted zeroes of all the models (52). This is likely due to its overprediction of horizontal plume spreading; a wide plume is more likely to cover many grid points leading to a fewer number of predicted zeroes at receptor

Table 5-1. Frequency Distribution of Predicted and Observed Concentrations for the Two Oklahoma Cases (Based on Predicted/Observed Pairings in Space and Time) in parts per 10^{15}.

	MESOPUFF		MESOPLUME		MSPUFF		MESOPUFF II		ARRPA		RTM-II		RADM	
	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED
= 0	47	94	39	114	53	109	26	88	31	80	51	52	15	97
(0 - 100)	122	56	122	33	122	53	122	46	122	60	122	89	122	23
(100 - 200)	13	18	13	17	13	14	13	14	13	9	13	25	13	4
(200 - 300)	11	10	11	6	11	6	11	4	11	9	11	13	11	6
(300 - 400)	4	10	4	5	4	4	4	7	4	8	4	7	4	5
(400 - 500)	2	10	2	7	2	4	2	3	2	7	2	2	2	2
(500 - 600)	2	0	2	5	2	5	2	6	2	3	2	4	2	1
(600 - 700)	2	1	2	3	2	0	2	3	2	1	2	8	2	0
(700 - 700)	1	1	1	2	1	4	1	5	1	3	1	6	1	5
(800 - 900)	2	0	2	3	2	2	2	1	2	1	2	2	2	3
(900 - 1000)	3	0	3	1	3	2	3	4	3	4	3	0	3	1
(1000 - 1500)	5	5	5	3	5	3	5	3	5	7	5	16	5	9
(1500 - 2000)	1	4	1	1	1	6	1	7	1	1	1	4	1	7
(2000 - 3000)	6	7	6	5	6	6	6	6	6	4	6	0	6	7
(3000 - 4000)	1	3	1	2	1	1	1	4	1	5	1	0	1	3
(4000 - 5000)	1	5	1	5	1	1	1	2	1	3	1	0	1	7
(5000 - 6000)	1	0	1	2	1	1	1	0	1	0	1	0	1	5
(6000 - 7000)	0	0	0	1	0	1	0	0	0	0	0	0	0	0
(7000 - 8000)	0	0	0	0	0	0	0	0	0	1	0	0	0	2
(8000 - 9000)	0	0	0	0	0	0	0	0	0	1	0	0	0	1
(9000 - 10000)	0	0	0	0	0	0	0	0	0	0	0	0	0	1
(10000 - 15000)	0	0	0	1	0	6	0	0	0	1	0	0	0	3
(15000 - 20000)	0	0	0	0	0	1	0	0	0	0	0	0	0	0
(20000 - 25000)	0	0	0	0	0	1	0	0	0	0	0	0	0	0
(25000 - 30000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
> 30000	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Totals:	224	224	216	216	230	230	203	203	208	208	228	228	192	192

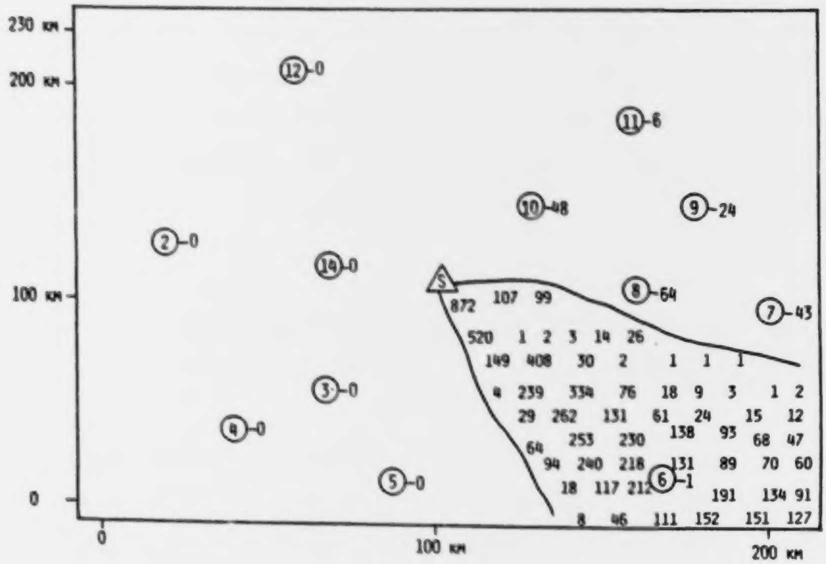
locations. Note also that MSPUFF and MESOPLUME have the largest number of observed zeroes. An examination of the patterns reveals that MSPUFF generally underpredicts horizontal spreading and MESOPLUME tends to have a combination of less spreading and less directional accuracy than the other models at the 100 and 600 km arcs.

5.3. COMPARISON OF PREDICTED CONCENTRATION ISOPLETHS WITH DATA AT THE SAVANNAH RIVER PLANT

The small number and large separation distances between sampling locations in the Savannah River Plant data base give poor spatial concentration gradient resolution, making such contour plots less useful for comparing model predictions with observed data. The added information that might be obtained from such isopleth plots did not justify the effort needed to prepare the large number of plots for the Savannah River Plant data base. Instead, computer printouts of average concentrations at each grid point were prepared for each model for each sampling period, with observed values printed below the grid points nearest them. In Section 5-4, the 65 predicted patterns for each model are broken down into 8 categories giving different relationships between the predicted and observed patterns. In this section, detailed plume outline graphs were drawn in the spirit of the isopleth plots for two 10-hour time periods represented in Figures 5-13 through 5-20. The graph and axis labels provide the same information as do the isopleth plots for Oklahoma. Each figure presents measured 10-hr averaged concentrations (above background) at each sampler (in pCi/m^3) along with values of the predicted plume concentrations within the outlined plume. A number within a circle represents the identification number (2 through 13) of a fixed sampler. Case 4B (Figures 5-13 to 5-16) illustrates one of the two common problems revealed with the model predictions for the SRP cases. All of the models appear to have a rotation angle error of the predicted plume as compared with the data. The models that use MESOPAC have a definite clockwise rotation to their predicted plumes. The models that use MESOPAC II exhibit the same rotation error, but by a smaller rotation angle.

Figures 5-17 through 5-20 illustrate the second problem found upon examining the patterns from the SRP case runs. This problem is an

MESOPUFF -- CASE 4B NOV 18 (2200 HR) TO NOV 19 (0800 HR)



MESOPLUME -- CASE 4B NOV 18 (2200 HR) TO NOV 19 (0800 HR)

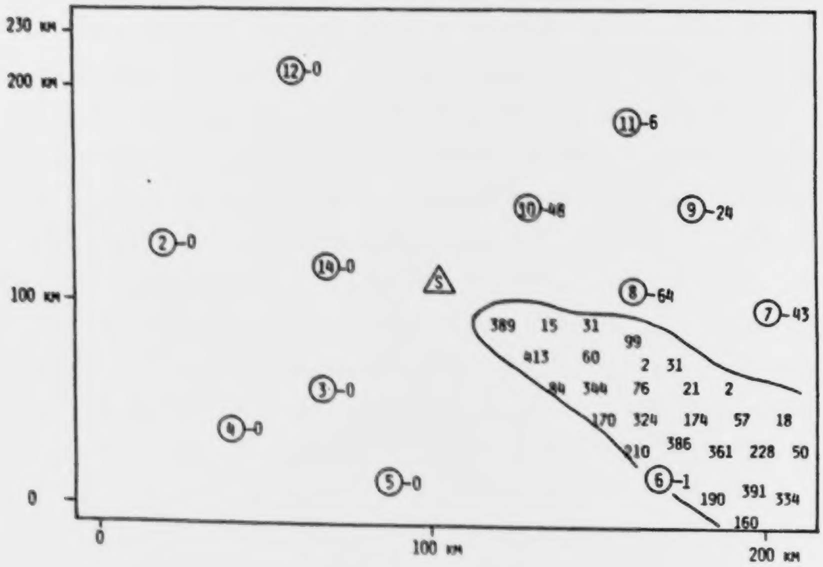
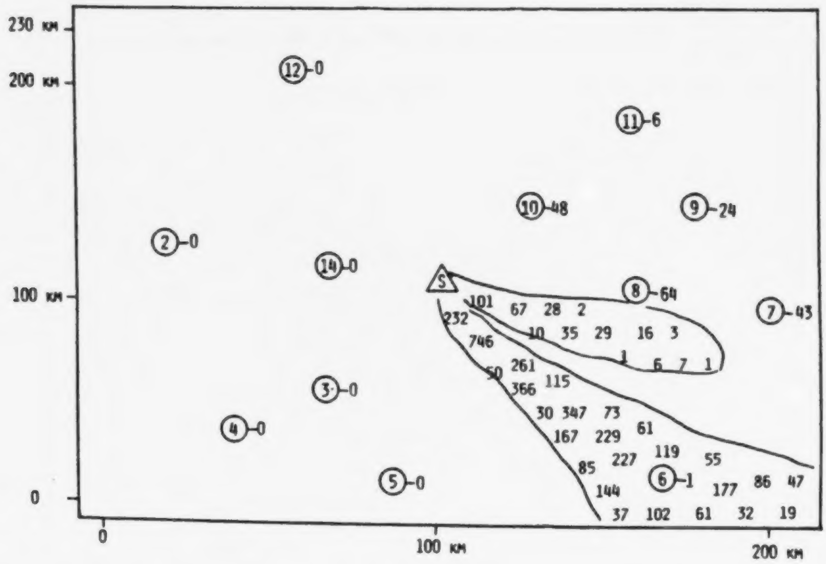


Figure 5-13. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... (top) MESOPUFF predictions, (bottom) MESOPLUME predictions.


MSPUFF -- CASE 4B NOV 18 (2200 HR) TO NOV 19 (0800 HR)

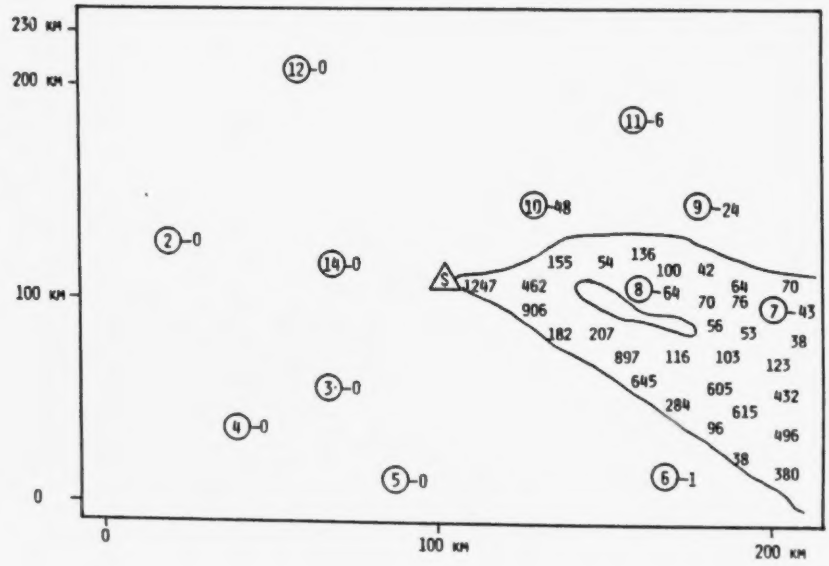

MESOPUFF II -- CASE 4B NOV 18 (2200 HR) TO NOV 19 (0800 HR)

Figure 5-14. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ...
(top) MSPUFF predictions, (bottom) MESOPUFF II predictions.

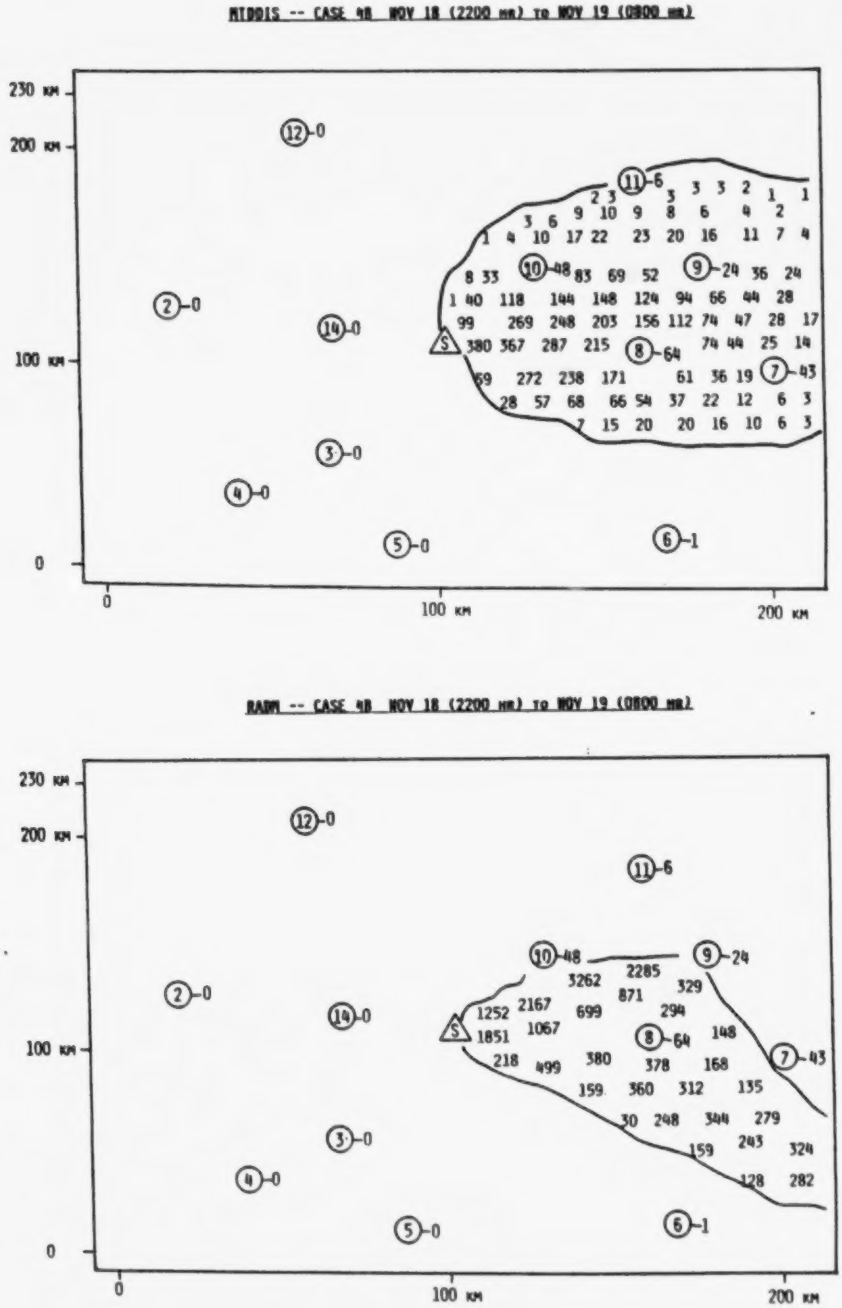


Figure 5-15. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... (top) MTDDIS predictions, (bottom) RADM predictions.

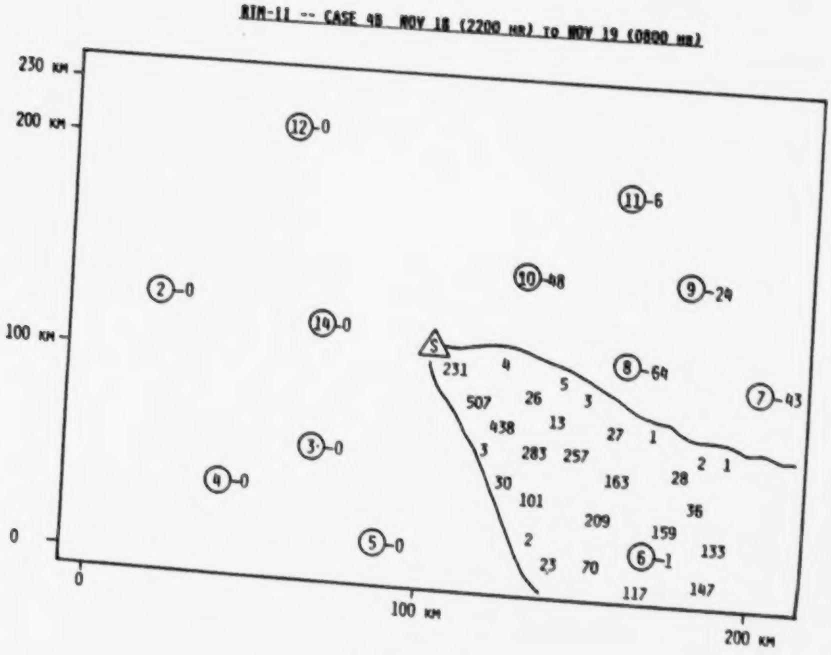
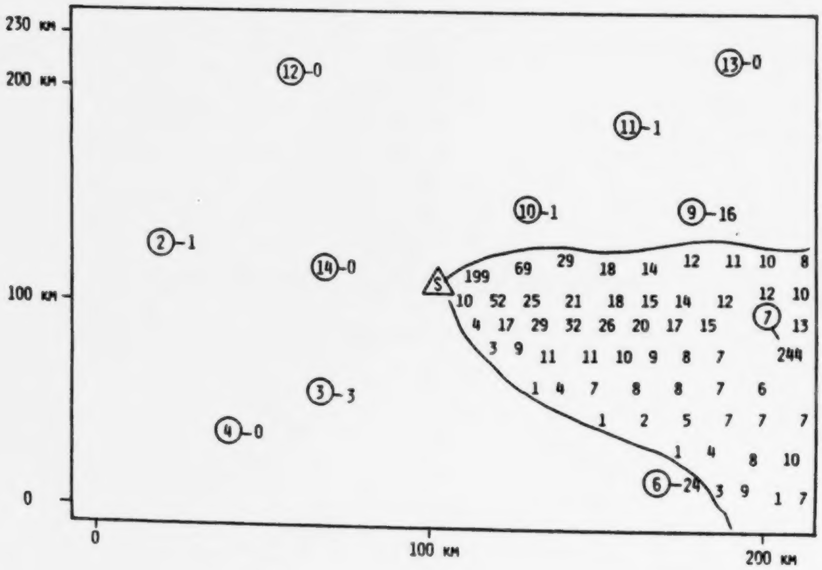


Figure 5-16. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of November 18-19, 1976 (2200 to 0800 GMT) ... RTM-II predictions.

MESOPUFF -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)



MESOPLUME -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)

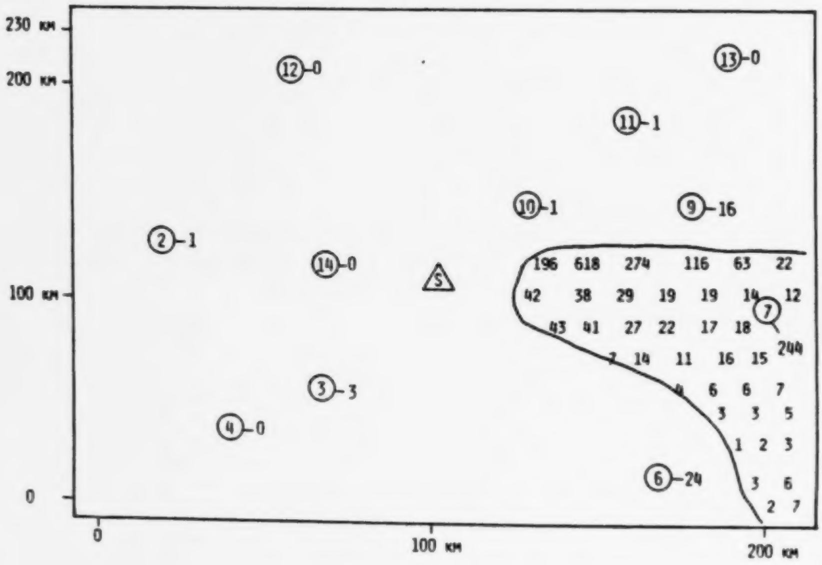
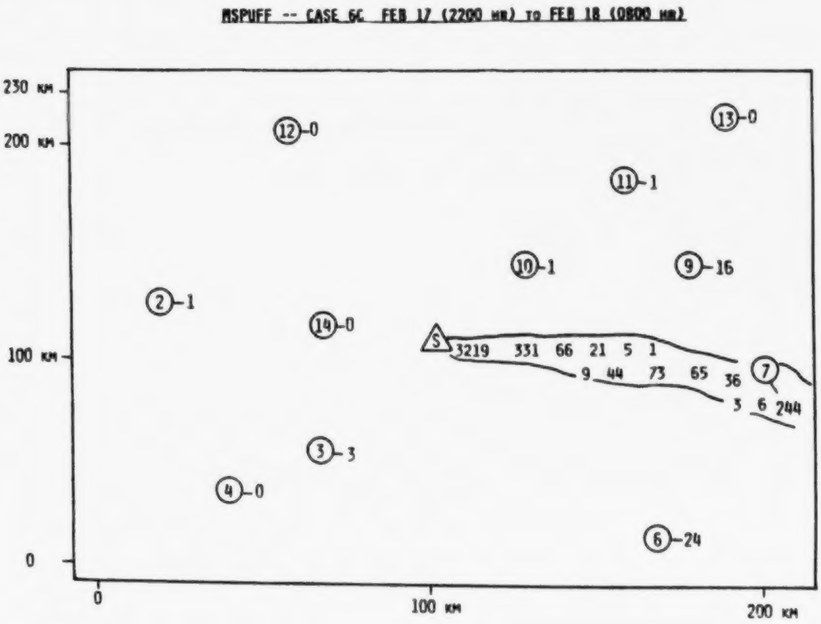


Figure 5-17. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... (top) MESOPUFF predictions, (bottom) MESOPLUME predictions.


MSPUFF -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)

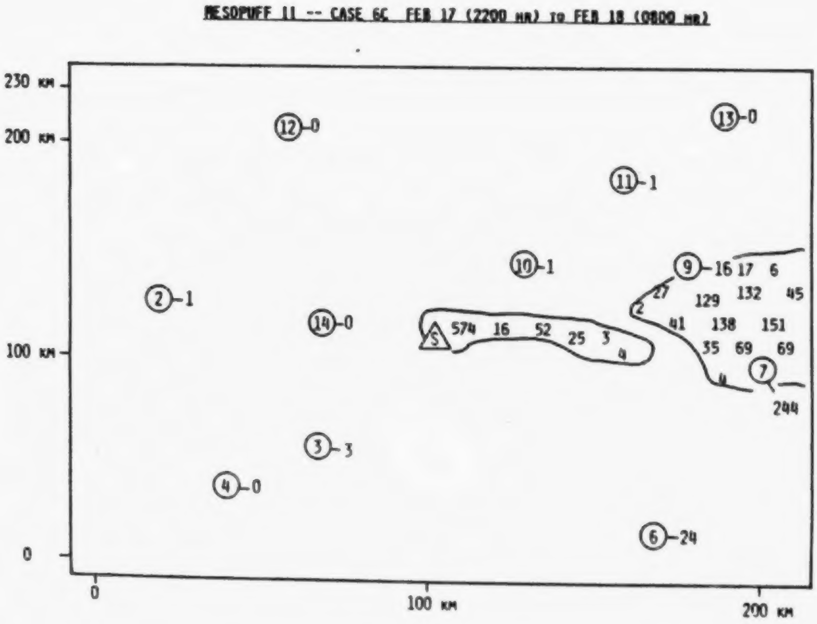
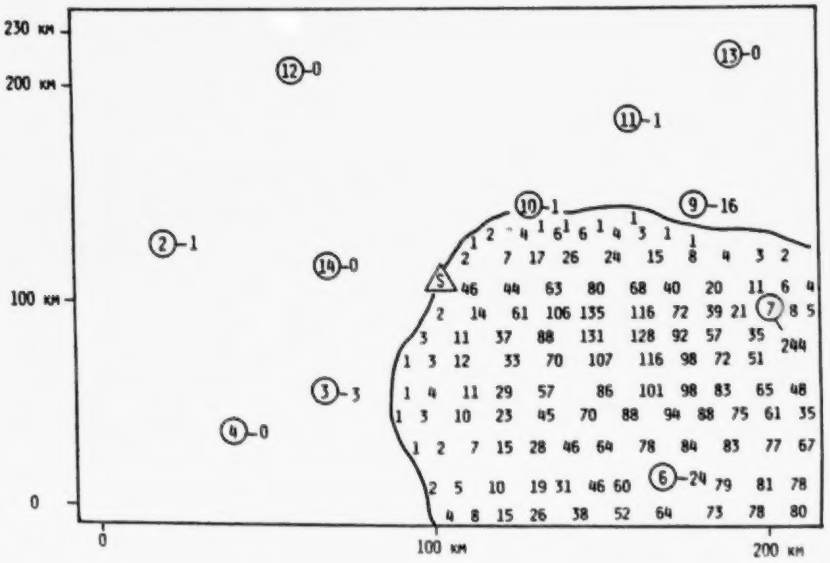

MESOPUFF II -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)

Figure 5-18. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ...
(top) MSPUFF predictions, (bottom) MESOPUFF II predictions.

MTDDIS -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)



RADM -- CASE 6C FEB 17 (2200 HR) TO FEB 18 (0800 HR)

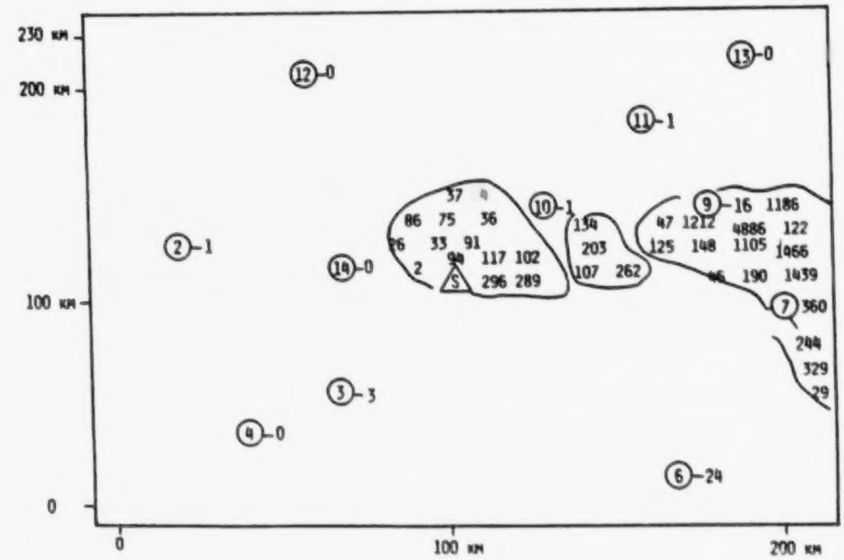


Figure 5-19. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... (top) MTDDIS predictions, (bottom) RADM predictions.

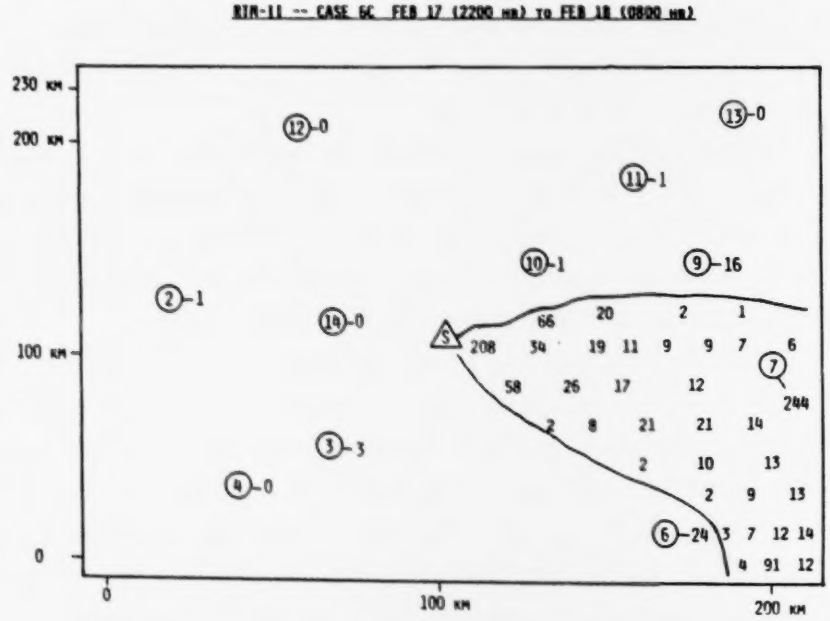


Figure 5-20. Comparison of 10-hour averages of predicted plume and observed data (in pCi/m^3) for Savannah River Plant experiment of February 17, 1977 (2200 to 0800 GMT) ... RTM-II predictions.

underprediction of the horizontal spreading of the plume by the models. Recall here that the RTM-II model which overpredicted spreading in the Oklahoma cases now underpredicts spreading. Two factors are relevant here for RTM-II. First, at SRP, the MESOPAC meteorological preprocessor is used rather than MESOPAC II. Second, and more importantly, the dispersion constant η and the minimum allowed value of the horizontal diffusivity were chosen differently in each case. RTM-II model predictions are very sensitive to the values used for these parameters. More discussion of these parameters and the RTM-II model is presented in Section 5.5.

Table 5-2 presents in similar format to Table 5-1 a breakdown of predicted and observed concentrations into a number of ranges. Whereas all models overpredicted peak values for Oklahoma, there is only a slight overprediction (in terms of numbers of large concentration values) for the SRP cases. Note that the MTDDIS and RADM models have a definite tendency to overpredict peak values and that RADM has one very large prediction greater than 30,000 pCi/m^3. Examination of the distribution of 10-hr concentration predictions greater than 200 pCi/m^3 for MTDDIS reveals a large number of predictions between 200-700 pCi/m^3, but few above 700 pCi/m^3. This may be explained by the large spreading observed with MTDDIS (as compared with other models); i.e., the predicted plume covers more samplers than other models leading to a larger number of predictions in each frequency category. Also, relatively few very large 10-hr concentrations are predicted with MTDDIS. An examination of the table (and the set of predicted ground patterns of the models) also reveals the following: MSPUFF and MESOPLUME have a combination of the least spreading and the least correct directional orientation of the models. One may verify this by noting the large number of predicted zeroes (214 for MESOPLUME and 206 for MSPUFF) in the first row of Table 5-2. The possible causes of the model/data discrepancies are discussed in Section 5.5.

5.4. PATTERN COMPARISON METHOD OF MODEL EVALUATION (OKLAHOMA ONLY)

It is interesting to evaluate the accuracy of the ground-level patterns predicted by the models as compared with the ground-level patterns observed in the data. It is clear from Figures 5-1 through 5-20 that the predicted and

Table 5-2. Frequency Distribution of Predicted and Observed Concentrations for the 15 Savannah River Plant Cases (Based on Predicted/Observed Pairings in Space and Time) in pCi/m^3.

	MESOPUFF		MESOPLUME		MSPUFF		MESOPUFF II		ARRPA		RTM-II		RADM	
	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED	OBS	PRED
= 0	27	193	22	214	52	206	57	133	167	46	46	146	9	196
(0 - 100)	240	84	240	54	240	89	240	154	240	328	240	138	240	35
(100 - 200)	21	8	21	12	21	18	21	30	21	34	21	15	21	23
(200 - 300)	8	9	8	1	8	6	8	6	8	15	8	10	8	13
(300 - 400)	4	2	4	9	4	1	4	4	4	8	4	5	4	5
(400 - 500)	2	2	2	2	2	4	2	4	2	1	2	3	2	3
(500 - 600)	0	2	2	2	2	3	2	2	2	3	2	4	2	2
(600 - 700)	3	1	3	1	3	1	3	1	3	7	3	3	3	0
(700 - 700)	1	3	1	2	1	1	1	3	1	3	1	2	1	1
(800 - 900)	0	0	0	2	0	1	0	0	0	1	0	1	0	1
(900 - 1000)	0	0	0	1	0	0	0	0	0	0	0	0	0	0
(1000 - 1300)	3	5	3	3	3	3	3	2	3	2	3	2	3	8
(2500 - 2000)	2	1	2	1	2	0	2	1	2	2	2	1	2	2
(2000 - 3000)	1	1	1	2	1	3	1	2	1	2	1	1	1	3
(3000 - 4000)	0	0	0	1	0	1	0	0	0	0	0	0	0	0
(4000 - 5000)	0	0	0	0	0	0	0	0	0	0	0	0	0	1
(5000 - 6000)	0	1	0	0	0	0	0	0	0	0	0	0	0	0
(6000 - 7000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(7000 - 8000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(8000 - 9000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(9000 - 10000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(10000 - 15000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(15000 - 20000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(20000 - 25000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(25000 - 30000)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
> 30000	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals:	224	224	216	216	230	230	203	203	208	208	228	228	192	192

observed plumes are generally offset spatially from one another. This feature is complicated by the fact that there is an offset in time as well, as the main bulk of the predicted plume arrives earlier or later at a given arc than the observed plume. Furthermore, it is theoretically possible for a model to predict excellent concentration patterns at the ground (as compared to similar patterns for the data), yet for the predicted and observed patterns to be simply offset in space and time from each other. The positioning problem in space and time that exists creates difficulties in interpreting the AMS statistics (Section 4.2) since those statistics provide an evaluation of the models based only on pointwise comparisons of model predictions and data. The AMS statistics do not separate out dispersion errors from directional errors.

In order to provide a simple estimate of the accuracy of the predicted patterns, an analysis was performed of the model predictions as they were available for the Oklahoma cases. In order to provide the maximum spatial resolution, the 34 x 34 set of gridded concentrations were not used because the interval between grid points was 20 km. Instead, the predicted hourly concentration values for the nongridded receptors were utilized, since they were generally much closer together, especially on the 100 km arc. The analysis was not performed for the SRP data sets since it was found that no clearly-outlined pattern in the observed data could readily be identified for most 10-hour averaging periods because of the sparcity of receptors and their wide spacing. However, a simplified yet quantitative pattern analysis of the SRP data cases will be presented in the next section.

The pattern comparison at Oklahoma was carried out by analyzing the average plume ground concentration pattern during the period of peak observed concentration as the plume passed each of the following arcs:

* 100 km arc for the July 8, 1980 experiment at Oklahoma
* 600 km arc for the July 8, 1980 experiment at Oklahoma
* 100 km arc for the July 11, 1980 experiment at Oklahoma

The peak periods were: 2230-2315 GMT for the 100 km arc of the July 8 experiment, 0800-1100 GMT for the 600 km arc of the July 8 experiment and 2330-0015 GMT for the 100 km arc of the July 11 experiment. Note that the observed data has averaging periods of either 45 minutes or 3 hours, whereas for the predicted values at the nongridded receptors, the hourly average values were used to provide greater temporal resolution for plume passage. In the case of the 45 minute averaged data at the 100 km arc, the difference is minimal; however, for the 3 hour averaged data at the 600 km arc, the observed arrival time was taken as the start of the period. Therefore, the predicted arrival times for the 600 km arc should be considered as uncertain within 1.5 hours. Fixing these time periods of peak concentrations, three variables were extracted from the model predictions and the observed data. The first two (centerline azimuth and transport time) describe the location of the plume, whereas the third variable (plume width) describes the spreading about that plume location. These variables are described in more detail below:

* centerline azimuth....the angular direction measured clockwise from north where the peak of the predicted or observed concentrations passes the 100 or 600 km arc of interest. The peak in the predicted plume usually occurred during a different time averaging period than for the observed data.

* transport time........the time from initial release to the time that the first traces of the plume passed the arc of interest.

* plume width...........the width of the plume (edge to edge) as the peak concentration passes the 100 or 600 km arc. Each predicted plume has a different transport time (not shown) for this peak to reach the arc of interest.

Tables 5-3 to 5-5 present the results of the these pattern comparisons for the 7 applicable models. (MTDDIS was not run for the Oklahoma data base.) Conclusions that may be drawn from an examination of these tables are:

(a) At the 600 km arc, the observed plume widths were underestimated by all the models. The observed width of the plume at the 100 km arc could not be estimated accurately due to horizontal spreading beyond the samplers. (The extreme sampler to the east or to the west showed high observed concentrations during the peak period.) The underestimation of predicted plume widths is consistent with the conclusions in Section 5.2 and with the general overprediction in plume concentrations found with the models. Considering that small predicted values were being compared with observed values near background in determining plume width, detailed width comparisons

Table 5-3. Pattern Comparison Results for the 100 km Arc ... July 8, 1980; Oklahoma Experiment.

Model/Obs.	Centerline Azimuth	Transport Time	Plume Width
Observed	358 degrees	3.0 hours	> 50 km
MESOPUFF	000	2.5	35
MESOPLUME	004	1.5	40
MSPUFF	356	2.0	25
MESOPUFF II	020	3.0	35
ARRPA	005	3.0	60
RADM	020	3.0	40
RTM-II	015	4.0	75

Table 5-4. Pattern Comparison Results for the 600 km Arc ... July 8, 1980; Oklahoma Experiment.

Model/Obs.	Centerline Azimuth	Transport Time	Plume Width
Observed	010 degrees	< 13.0 hours	200 km
MESOPUFF	024	15.0	140
MESOPLUME	024	16.0	140
MSPUFF	017	13.0	100
MESOPUFF II	011	20.0	160
ARRPA	019	15.0	100
RADM	024	22.0	50
RTM-II	024	21.0	180

Table 5-5. Pattern Comparison Results for the 100 km Arc ... July 11, 1980; Oklahoma Experiment.

Model/Obs.	Centerline Azimuth	Transport Time	Plume Width
Observed	019 degrees	3.75 hours	> 60 km
MESOPUFF	002	2.5	40
MESOPLUME	002	2.5	45
MSPUFF	002	3.5	35
MESOPUFF II	002	4.0	60
ARRPA	015	4.0	60
RADM	002	4.0	40
RTM-II	352	4.0	60

among models here are not justified. On the other hand, the models that predict plumes that are generally much too narrow (such as MSPUFF and RADM) could be distinguished during at least two of the three periods.

(b) The centerline azimuth results indicate that the accuracy of the direction of the peak concentration as it passes the arcs is quite variable between arcs and experiments. Direction predictions for MESOPUFF II, RADM, and RTM-II were significantly shifted to the east for the 100 km arc on the July 8 experiment, and showed the same shift during passage through the 600 km arc.

(c) The transport time comparisons indicate that the predicted plumes of three models reach the 100 km arc more rapidly than the observed plume. However, for the 600 km arc, all predicted plumes lag the observed plume.

(d) No model shows a consistently superior performance (as compared with the other models) with respect to the spreading and location variables presented in Tables 5-3 through 5-5. More accurate and more definitive pattern comparison results could be obtained in an experiment where sampler arcs are always wide enough to include all of the observed plume, where receptors are operational during the entire episode of plume passage, and where sampling utilizes shorter and more frequent averaging times. The models would also show more precise arrival times if they were modified to provide shorter-term averages than every hour.

(e) Although MESOPUFF II, RADM and RTM-II share the same meteorological preprocessor, their centerline azimuth, travel time and plume width values are not identical in these tables. MESOPUFF II differs more significantly from the other two models, because it alone uses the upper-level wind field generated by MESOPAC II in transporting the plume. The other two models rely solely on the wind field averaged over the mixing layer. The small differences between their predicted values for these parameters stem primarily from the differences in how they handle exchange between the upper layer and the mixing layer (RADM has only one layer), and in how they compute surface concentration values from elevated plume material.

5.5. SUMMARY OF DIAGNOSTIC REVIEW OF MODEL/DATA COMPARISONS

In this section, the predictive performance of each of the eight models is related to the theoretical formulation of the model. The primary tools for interpretation, in addition to the statistical measures already presented, are:

(a) the contour plots of averaged concentration for each of the 21 time averaging periods of the two Oklahoma experiments (as shown in Figures 5-1 to 5-12 and Figures E-12 to E-95), and

(b) the printed panels of 10-hour-average gridded concentrations and observed values for the 65 time averaging periods for the Savannah River Plant experiment (as shown in Figures 5-13 to 5-20).

The graphical comparisons cited above for Oklahoma (not available for MTDDIS) have good spatial resolution at the two arcs of samplers, and good time resolution for passage of the plume through these arcs. Those graphs also allow evaluation of model performance over distances as large as 600 km. However, those releases took place under fairly steady summertime conditions, and model performance under a variety of meteorological conditions cannot be inferred from that data. The graphical comparisons for the SRP data sets (not available for ARRPA) offer much less spatial resolution, with only 13 samplers located within about 150 km of the source. The SRP comparisons have much coarser time resolution, since the averaging periods are all at least 10 hours; however, those comparisons include approximately equal numbers of nighttime and daytime averaging periods, and includes cases from all four seasons of the year.

Three supplementary tables have been prepared to provide additional quantitative measures of model predictive performance at each of the two sites. Table 5-6 lists the maximum concentration that each model predicts (in parts/10^{15}) over the whole prediction grid for each of the 21 averaging periods during the Oklahoma experiment. For each of the averaging periods, a ranking of the seven models was also prepared, from 1 (lowest maximum concentration) to 7 (highest maximum concentration). The small subtable at the bottom of Table 5-6 shows the average ranking for each model over the 21 Oklahoma averaging periods. Table 5-7 presents the same analysis of model predictions in pCi/m^3 for the 65 Savannah River Plant data cases (with MTDDIS predictions replacing those of ARRPA).

The third table (Table 5-8) requires more description. Each of the 65 Savannah River Plant cases for each model is classified with respect to two major features. First, each subcase was counted in one of 4 categories that indicates roughly the comparison of the predicted plume trajectory with the observed one. Some subjective judgment was required in assigning predicted plume patterns to directional categories, because sometimes only one or two samplers had observed values above background. Other times all of the observations were within several pCi/m^3 of the background of 16, and one or more of those samplers could have been recording fluctuations in the background concentration. Secondly, each case was counted in one of 3 categories that indicates whether the plume spread too much, too little or about the right amount (implied by category (1) with no entry in (7) or (8).) Each of these three tables will be referred to in the summary discussion below of each model whenever the table is useful in drawing inferences concerning the degree of validity of a model's theoretical formulation.

Issues to receive special mention are (a) the adequacy of the wind field and the gridded mixing heights predicted by the meteorological preprocessor, and (b) the accuracy of the spreading rate implied by the model's formulation of plume dispersion. Also to receive some mention, where possible, is the method each model uses to compute ground concentrations from its representation of the three-dimensional spatial distribution of plume material.

Table 5-6. Summary of Peak Predicted Concentrations for Oklahoma Data Sets, July 1980 (parts/10^{15}).

Day/Time(GMT)	MESOPUFF	MESOPLUME	MSPUFF	MESOPUFF II	ARRPA	RADM	RTM-II
08 2100-2145	9185	4807	180365	822884	7139	147472	4576
08 2145-2230	8014	3904	124975	274294	26494	63413	4040
08 2230-2315	6272	4173	64854	4375	26494	28028	4845
08 2315-0000	5180	4217	20190	7609	7713	29047	5262
09 0000-0045	4079	6574	12558	3447	8426	17227	2876
09 0045-0130	2804	4918	9635	3222	5531	9373	1790
09 0130-0215	2860	4578	7798	2115	5695	7074	1206
09 0800-1100	1017	1147	924	1575	3072	9831	88
09 1100-1400	869	976	896	825	2462	4974	44
09 1400-1700	767	956	626	387	2258	1026	12
09 1700-2000	662	941	0	421	217	2232	3
09 2000-2300	546	961	0	419	0	1296	1
09 2300-0200	335	1033	0	214	0	418	0
11 2200-2245	542	653	5367	812	935	4105	748
11 2245-2330	438	566	4538	838	945	2154	673
11 2330-0015	439	557	3302	997	694	2310	637
11 0015-0100	407	780	3110	1149	773	2442	709
11 0100-1045	418	831	1323	762	725	1836	315
11 0145-0230	355	572	1238	677	909	1730	165
11 0230-0315	296	489	1198	470	845	1343	125
11 0315-0400	318	569	1053	478	723	1295	111

Average rank among models of maximum concentration by case:
(1 = lowest max. conc., 7 = highest max. conc.)

	Average
RTM-II	1.86
MESOPUFF	2.76
MESOPLUME	3.67
MESOPUFF II	3.81
ARRPA	4.38
MSPUFF	5.24
RADM	6.38

Table 5-7. Summary of Peak Predicted Concentrations for Savannah River Plant Data Sets, pCi/m^3.

Case	MESOPUFF	MESOPLUME	MSPUFF	MESOPUFF II	MTDDIS	RADM	RTM-II
1 --	1740	2241	4049	12820	28581	8691	1940
2A --	19209	10604	7451	2398	3694	16875	3595
2B --	1609	322	183	747	959	16185	281
2C --	12936	11581	17269	3867	51	11250	5446
3A --	1675	936	891	39425	20874	3163	558
3B --	1467	1227	735	1660	3299	5276	455
4A --	240	236	509	21040	3467	623	254
4B --	872	391	746	8253	380	3522	508
4C --	1747	374	613	26014	2753	2924	319
4D --	0	0	0	0	103	0	0
5A --	3978	1995	1487	23495	3077	2367	964
5B --	10509	13622	1201	10606	245	13557	1921
5C --	3373	1693	3348	26222	1883	32283	1470
5D --	4794	1596	3446	20559	217	4658	540
6A --	9753	2657	7186	19999	154	8958	2193
6B --	1949	1390	3757	45342	1599	10408	1310
6C --	199	618	3219	574	135	4886	208
6D --	2937	2042	51641	7830	5159	11531	1533
6E --	292	359	12513	17398	1147	3058	280
7A --	399	339	251	2078	552	6605	245
7B --	0	0	0	0	34	0	0
8A --	141	160	366	1756	0	2950	180
8B --	3551	586	2564	56844	3147	5708	801
8C --	288	258	886	7700	3818	2075	187
8D --	1053	566	2273	46643	6466	2290	222
8E --	1607	198	477	5378	1446	1352	174
8F --	98	66	351	1355	1262	700	57
8G --	0	0	0	0	26	0	0
9A --	12771	3782	1564	49340	4051	12969	3885
9B --	832	4824	40	1503	339	3102	606
9C --	303	1514	0	412	298	3294	481
9D --	7383	144	900	6153	3370	9109	1447
9E --	4532	6847	1914	31018	2879	6764	484
9F --	247	238	183	1096	440	4961	178
9G --	19	24	101	759	83	1409	49
9H --	3	0	1644	6839	907	3263	75
9I --	2975	1331	2465	48820	1537	4508	724
9J --	1279	2230	20004	4214	1968	4327	1198
10A --	5066	13141	20590	10507	27851	35005	1328
10B --	6957	1813	5275	16676	4124	30533	5055
10C --	3877	5686	17688	79640	11369	13089	4946

Table 5-7. (continued) Summary of Peak Predicted Concentrations for Savannah River Plant Data Sets, pCi/m^3

Case	MESOPUFF	MESOPLUME	MSPUFF	MESOPUFF II	MTDDIS	RADM	RTM-II
10D --	6825	4159	11311	38669	1820	10326	1048
10E --	160	713	666	6459	207	5925	189
10F --	26188	2257	5724	67818	66666	22897	4363
10G --	424	987	1578	12682	1968	2231	620
10H --	7054	3304	7468	82569	17589	13269	1869
10I --	822	753	1429	19341	1763	4346	615
10J --	383	242	171	19465	3511	2954	138
11A --	369	404	123	4928	244	7356	139
11B --	499	646	809	16072	4440	1902	587
11C --	3409	3392	8251	38785	65168	6720	1750
11D --	357	270	494	9069	8938	1232	306
12A --	3210	2579	3115	53283	3201	5312	920
12B --	3547	7844	18072	15054	10123	17664	4125
13A --	5562	3856	554	18265	4787	6081	588
13B --	162	220	0	60	625	1273	96
14A --	0	0	0	0	0	0	0
14B --	51	1114	118	0	16	904	27
14C --	74518	13106	10631	74888	16859	37242	3604
14D --	5293	1656	11802	37869	1357	2904	1177
14E --	315	923	343	871	388	1306	258
15A --	8554	7214	23214	79738	14603	10744	1378
15B --	1263	2025	23003	48624	2965	3565	1400
15C --	344	284	690	1547	754	2813	251
15D --	119	141	0	193	315	2531	42

Average rank among models of maximum concentration by case:
(1 = lowest max. conc., 7 = highest max. conc.)

	Average
RTM-II	1.80
MESOPLUME	3.05
MESOPUFF	3.56
MSPUFF	3.72
MTDDIS	4.11
RADM	5.79
MESOPUFF II	5.97

Table 5-8. Classification of Model Prediction Types for the Savannah River Plant.

Predicted Plume	MESOPUFF	MESOPLUME	MSPUFF	MESOPUFF II	MTDDIS	RADM	RTM-II
Categories related to plume trajectory:							
1 - Right direction	30	25	15	28	47	28	31
2 - Rotated clockwise	14	18	24	20	5	20	13
3 - Rotated counter-clockwise	9	9	9	10	9	11	9
4 - Rotated appx. 180 degrees	6	6	7	1	1	1	6
5 - Complex, confused	2	1	1	1	1	1	2
6 - Pred. all 0; Observed not	4	6	9	5	2	4	4
Categories related to plume diffusion rate:							
7 - Narrow	26	28	24	10	9	22	12
8 - Wide	0	0	1	8	24	0	1

NOTES

(1) The predicted plume direction was correct within about 10-20 deg. as compared with the observed plume.
(2) The predicted plume was rotated clockwise (as viewed from above) at least 20-45 deg. from the observed plume.
(3) The predicted plume was rotated counterclockwise (as viewed from above by at least 20-45 deg. from the observed plume.
(4) The predicted plume direction appeared to lie about 180 deg. from that indicated by the available data.
(5) No simple relationship existed between the predicted and observed plumes, but there was significant disagreement between them.
(6) The predicted plume concentrations were zero at all grid points, but there was at least one observation above background.
(7) The predicted plume was narrower than the observed plume.
(8) The predicted plume was wider than the observed plume.

5.5.1. MESOPUFF

In the evaluation of the performance of MESOPUFF, it is useful to try to separate the causes of model/data discrepancies between the meteorological preprocessor MESOPAC and the plume dispersion model MESOPUFF. The wind field, mixing heights, and stability classes are provided by the meteorological preprocessor MESOPAC. Plume growth and the calculation of ground concentrations are features of MESOPUFF. The separation of effects is impossible to carry out completely, since mixing heights and stabilities (computed in MESOPAC) influence plume spreading and concentration calculations (computed in MESOPUFF). One distinction, however, is that the wind field alone affects plume advection leading to the predicted trajectories. Obviously, conclusions concerning model/data discrepancies relating to the plume trajectory and times of arrival of the plume at fixed receptors are easier to make than conclusions relating to plume model performance (e.g., the rate of diffusion and the magnitude of predicted ground concentrations).

The discussion now focuses on the performance of MESOPAC in predicting the location and time of arrival of the predicted plume. The summary of the behavior of plume trajectories at SRP shown in Table 5-8 indicates that in nearly half of the cases (30 cases), the direction of plume travel is approximately correct. However, for a substantial fraction of the cases (29 cases) the plume was rotated either clockwise (14 times), counterclockwise (9 times) or advected in nearly the opposite direction from the data (6 times). The cause of the mild preference for clockwise rotation can be suggested. Since the wind field predicted by MESOPAC for SRP (and Oklahoma) is computed by averaging twice-daily rawinsonde winds over 1500 m above ground (and then interpolating in space and time), some of the effects of the Ekman spiral will be observed. The Ekman spiral involves rotation clockwise of the wind direction with elevation (as viewed from above) relative to the surface winds. It appears from examination of the meteorological data and results of the model runs that the surface wind direction is more appropriate for characterizing transport of the ground-level plume, even though the plume material is likely to be vertically mixed throughout the mixing layer roughly 20-50 km from the point of release. It is likely that wind direction shear causes mixing to occur in a skewed sense, following the Ekman spiral.

It is interesting that 6 cases occurred where the predicted plume was advected in nearly the opposite direction from the observed plume. The probable reason is the single-layer nature of the MESOPAC wind field. Under some meteorological conditions, the wind direction in an elevated layer is opposite to the surface wind, and a single-layer model like MESOPAC with 1500 m averaging will obtain a wind direction that includes the elevated layer. If the surface layer up to the inversion is shallow when the winds in the two lowest layers are opposite, the average wind direction will be nearly that of the elevated layer. Then the plume material that leads to ground concentrations will advect in the direction of the winds in the surface mixing layer, whereas the predicted plume will advect in the opposite direction. This interpretation is strengthened by noting from Table 5-8 that MESOPUFF II (whose wind field is provided by the two-layer MESOPAC II model) produces only one such case.

A second result of the simplifying assumptions used in the MESOPAC meteorological preprocessor is that during the nighttime averaging periods at SRP, new puff releases are likely to remain within two stack heights (124 m) of the ground, and a 1500 m vertical averaging of the horizontal winds will not correctly represent transport of these new puffs. Within MESOPAC, one could attempt a 1500 m averaged wind for daytime transport and a 150-m averaged wind for nighttime transport. MESOPAC, as formulated, requires an input wind value from the soundings before a mixing height is calculated. That procedure for averaging the winds over the local mixing height would lead to too low a wind speed at night for transport of the plume puffs already diffused aloft. The development of MESOPAC II was, in part, aimed at solving the inherent problems in defining a single layer wind field.

Examination of the Oklahoma results reveals that the plume passes through the 100 km arc at about the right location, but somewhat early. It passes the 600 km arc clockwise of the point where the samplers show maximum concentrations. The predicted plume then veers quite strongly to the east, possibly in response to an approaching front. The plume's arrival at 600 km is somewhat late (by at most 2-3 hours), probably due to the presence the night before of a strong elevated nocturnal jet (documented in Appendix E) which is only partially represented by a 1500-m averaged wind. This behavior tends to confirm the interpretation given above for the SRP results. The plume angular offset caused by using a wind field that is averaged over the Ekman spiral does not lead to a large lateral offset distance at the 100 km arc.

However, in general the angular error will lead to a large lateral offset distance at the 600 km arc.

The remarks which follow summarize the findings with respect to the plume model MESOPUFF. The pattern comparison analysis in Tables 5-3 to 5-5 indicates that the spread predicted by MESOPUFF is less than that of the observed plume at the 100 km arc, (as is the case for all of the models except RTM-II and ARRPA). However, examination of the entries in Table 5-6 for MESOPUFF shows that its predicted maximum concentrations at Oklahoma are the second lowest for the seven models applied to that data base.

The model apparently predicts insufficient lateral spreading as may be seen in the SRP model/data comparisons in Table 5-8. However, the maximum concentrations predicted by MESOPUFF usually rank in the middle of the maximum concentrations predicted by the other six models (an average rank order of 3.56 on a scale of 1 to 7), as demonstrated by Table 5-7. For very short-term averages (2-3 hours), such small spreading can be interpreted as inadequate treatment of small-scale processes, e.g. too small a value for σ_y. For 24-hour averages, too small a lateral spreading would most likely be due to wind shear and plume meander, i.e. inadequate resolution in the meteorological preprocessor. For 10-hour averages, as in the present case, a combination of both deficiencies is likely to be present.

Concerning small-scale processes, the use in MESOPUFF of the Turner curves out to 100 km does not provide sufficient spreading, especially under stable conditions. The Turner σ_y values are more applicable in the range up to 10 km and for flat level terrain. The transition point for the Heffter formulation should probably be no farther out than 10 km. Additional numerical experiments with these SRP data cases can better verify that point.

The results of the model/data comparisons for the first Oklahoma release tend to confirm, with higher spatial resolution, the observations of plume spreading evident in the SRP predictions. As the plume passes the first arc (100 km), its width and location are approximately correct, although predicted plume arrival is a little late. Very early plume dispersion should be adequately represented by the Turner curves. In MESOPUFF, the transition to the Heffter formulation occurs at 100 km, which is probably at too long a distance because large-scale processes have already become important; in any case, the effect of this late transition from the Turner curves to the Heffter formulation has not had significant effects on the predicted concentrations and spreading. As the plume arrives at the 600 km arc, it is still narrow by about 30%, due to the presence of early spreading rates that were too small.

5.5.2. MESOPLUME

The theoretical formulation of the MESOPLUME model is very similar to the MESOPUFF model, and the predictions of both models have many similarities. Both models use the same meteorological preprocessor, MESOPAC, with identical inputs and outputs. As a result, the trajectories and times of arrival of the plumes should be about the same for both models. In the treatment of plume dispersion, both models use the Turner curves to obtain stability-dependent Gaussian plume growth rates out to 100 km. Beyond 100 km, the formulation by Heffter is employed. The predictive differences that do emerge between the models must therefore derive from

(a) their different geometrical representations of the shape of the region that is spreading according to the Gaussian widths, and

(b) their formulas for predicting ground-level concentrations.

In MESOPLUME, the plume is viewed as a series of cylinder-like plume segments. In MESOPUFF, the plume is represented by spherical "puffs." In both cases the growth of lateral and vertical dimensions is governed by the Turner curves and the Heffter formulation, but in calculation of the ground concentrations, the geometrical differences in plume shape will lead to differences in model predictions. Thus, comparisons of the predictions of the two models sheds light on the differences that result from the two theoretical approaches to defining plume shape and the calculation of ground-level concentrations.

In spite of the numerous similarities in the models, an examination of the predictions for the two Oklahoma releases reveals that MESOPLUME and MESOPUFF have significantly different ground concentration predictions. For the first release (from 1900 to 2200 GMT on July 8, 1980), the maximum concentrations listed in Table 5-6 for MESOPLUME are nearly a factor of two less than the MESOPUFF predictions for the first 1.5 hours (from 2100 GMT).

However, the predictions of the two models become equal after 3 to 4 hours. Thereafter, MESOPLUME predicts higher maximum concentrations than does MESOPUFF. During the second release (from 1900 to 2200 GMT on July 11, 1980), the maximum concentrations predicted by MESOPLUME are higher than those predicted by MESOPUFF for the entire 6 hours after the release, beginning with a factor of about 1.2 and ending with a factor of 1.8.

At both arcs for the first release, and at the 100 km arc for the second release, the width of the MESOPLUME predicted plume, however, is equal to or greater than that of the MESOPUFF predicted plume. This behavior offers an interesting insight into the geometrical differences between these models. Unlike the SRP data sets, the Oklahoma data sets feature a ground-level release with the 100 km arc relatively close to the source. Apparently, the use of cylindrical plume segments in MESOPLUME as contrasted with the spherical plume puffs in MESOPUFF, and the use of slightly different methodologies for predicting ground concentrations from the elevated puffs, combine to provide an opposite trend when the material is released near the ground, particularly near the source. The larger concentrations listed for MESOPLUME in Table 5-6 for the first release (before passage through the 100 km arc) are, thus, coupled with a wider plume spread at the ground as compared with MESOPUFF. These differences come solely from differences in the method of calculating ground concentrations from plume parcel shape and not from differences in rates of plume diffusion. Although both models are examples of puff models, the methods used in MESOPUFF yield more accurate predictions in this study than the methods used in MESOPLUME.

The pattern analysis results in Tables 5-3 to 5-5 show that the times of arrival at the Oklahoma sampling arcs are identical to those of MESOPUFF within 1 hour. The 1 hour difference for the first release is due to different amounts of predicted spreading in the ground-level concentration pattern, because the underlying wind field is the same for both models. Inspection of any corresponding pair of concentration isopleth graphs for Oklahoma shows this differing amount of spread at ground-level. These differences in the spreading of the ground-level plume cannot be attributed to the use of different plume dispersion coefficient formulations. Rather, they come from differences in plume parcel geometry and differences in the method of computation of ground-level concentrations.

It is more difficult to draw definitive conclusions about the respective merits of the concentration averaging formulations of the two models for the SRP model/data comparisons because of the larger averaging periods for which these comparisons were made (10 hours) and due to the large separation between receptor locations. An examination of the data show that for those 10-hour averaging periods when receptors have concentrations above zero, it is usual that only one to three of them report the presence of the plume, and these points are not arranged in clearly defined arcs giving good spatial resolution. As mentioned previously, this blurs distinctions between the effects of the dispersion formulations and the effects resulting from the plume trajectory positioning. However, in general, one may say that MESOPLUME shows some tendency to overpredict the observed concentrations when the observed values are above background. This is evident both from the scatter plots of observed data versus predictions, the graphs of frequency of residual, the graphs of cumulative frequency distributions of observed and predicted values, and the graphs of average versus residual. The statistical results reveal this conclusion as well. There is only a moderate difference in predictions between the two models for the SRP data. More significant is the fact that MESOPUFF predicts nonzero concentrations 68 times when the observed concentration is nonzero, whereas MESOPLUME does so only 56 times. From that information, one may surmise that MESOPLUME leads to a smaller spreading at ground level than does the MESOPUFF model. Results in Table 5-8 support these conclusions. In 6 cases, MESOPLUME predicted zero concentrations at every grid point when at least one sampler showed the presence of tracer material, whereas MESOPUFF did so in only 4 cases. The MESOPLUME predicted plumes were in the right direction for 5 fewer cases than were the MESOPUFF predicted plumes. The directional error in the MESOPLUME predicted plume was a clockwise rotation from the observed plume in 4 of the 5 cases for which the plume predicted by MESOPUFF was in the right direction.

The conclusion that the MESOPLUME predicted ground patterns were narrower than those of MESOPUFF for the SRP experiment should be contrasted with the conclusion for the Oklahoma experiment that MESOPLUME seemed to have broader ground concentration patterns. By mass conservation, a narrower predicted plume should contain higher maximum concentrations. However, an examination of Table 5-7 shows that MESOPLUME predicts lower peak concentrations on the average than does MESOPUFF. This difference, opposite to what is expected with plume material uniformly distributed in the mixing layer, must be a result of the different methods for computing ground concentrations from elevated parcels. For the Oklahoma data sets, the effects of the cylindrical and spherical puffs are being compared with the centers of both types of puff relatively near the ground, whereas for the Savannah River Plant data sets they are being compared with the centers of both types of puff elevated by at least 62 m. Apparently, the effects of these differences in puff geometry on ground concentrations are greatest when the puff centers are near the ground. In summary, for the Oklahoma experiment, the plume at ground level predicted by MESOPLUME is wider and has larger concentrations than the plume predicted by MESOPUFF. On the contrary, the MESOPLUME predicted plumes for the Savannah River Plant data base were narrower and had lower concentrations than the plumes predicted by MESOPUFF.

5.5.3. MSPUFF

It might be expected that the similarity of the formulation of MSPUFF to that of MESOPUFF would lead to very similar predictive behaviors for both models. However, significant differences emerge from a comparison of both models with the data. For this reason, the discussion of the predictive performance of MSPUFF as it relates to its formulation will focus on (a) the differences between the performance of MSPUFF and the performance of MESOPUFF, or (b) the predictive accuracy of MSPUFF with respect to observed data. An examination of Tables 5-3 to 5-8 yields the following performance characteristics of MSPUFF:

(a) The MSPUFF predicted ground concentration patterns often appear similar to the observed pattern in shape, but seem to be rotated about the source through some angular offset, and

(b) The ground-level concentrations are overpredicted with the width of the pattern usually narrower than the observed pattern.

Point (a) is evident from the centerline azimuth comparisons in Tables 5-3 and 5-4 for Oklahoma, and in categories (1) through (3) in Table 5-8 for the SRP data sets. At SRP the MSPUFF plume was in the right direction for only 15 of 65 cases, and was rotated clockwise in 24 cases. In the same tables, MESOPUFF shows a higher degree of directional accuracy. Since MSPACK, (the meteorological preprocessor for MSPUFF) employed 850 mb winds for its single-layer wind field, and the user option of averaging the winds over 1500 m was used for the MESOPAC runs, the wind directions predicted by MSPACK tended to be rotated clockwise from those of MESOPAC. This increases the clockwise rotational errors already present in MESOPAC due to the Ekman spiral effect. Model/data discrepancies involving the predicted trajectories may be explained largely by the choice of these 850 mb winds in MSPACK for both the Oklahoma and SRP data sets. For the SRP data sets, MSPUFF showed the greatest tendency of any model to predict a plume that has been rotated clockwise as compared to the data. This effect should be most pronounced during the nighttime hours, when convective mixing does not tend to eliminate the spiral. In only 15 of the 65 cases was the direction of the predicted plume judged to be correct, fewer than for any other model.

In support of point (b), Table 5-8 shows that plumes narrower than the observed plume were predicted in 24 cases. The tendency toward overprediction was documented in the statistical analysis of Section 4.2. Also, Table 5-6 indicates that peak ground concentrations predicted by MSPUFF for the Oklahoma data usually exceed those of MESOPUFF and MESOPLUME. Only RADM predicts larger peak concentrations. (Table 5-7 shows that peak ground concentrations predicted by MSPUFF for the SRP data exceed those of both MESOPUFF and MESOPLUME.) This physical behavior is currently an unresolved issue. Several possibilities have been examined, however. First, as discussed in Section 2, the Turner curves employed by the MSPUFF model have been altered slightly in the range less than 50 km in order that steady-state predictions of MSPUFF would agree with the MPTER model. This difference between dispersion coefficients can lead to some differences in plume concentration predictions for MSPUFF as compared to MESOPUFF. No extensive comparisons have been made of the differences, however. Second, the Heffter formulation for dispersion was used beyond 50 km from the source rather than 100 km as in MESOPUFF. This change should lead to lower concentrations in MSPUFF as compared to MESOPUFF beyond 50 km. The differences in dispersion coefficients are not, therefore, considered to be the main cause of the behavioral differences between the MESOPUFF and the MSPUFF predictions. By a process of elimination, it is suspected that the larger concentration predictions at ground level given by MSPUFF as compared to MESOPUFF are due to changes in the mixing height formulation. The two key changes made there in MSPUFF were:

(a) surface winds were used to compute the mechanical mixing height, and

(b) turbulent dissipation of convection was added to the computation of the convective mixing height in the Benkley-Schulman method.

It is not clear at this time what the precise effects of these changes were on the prediction of mixing heights during the diurnal cycle.

The pattern comparison results presented in Tables 5-3 to 5-5 also support these conclusions. The width of the MSPUFF predicted plume was consistently less than that of MESOPUFF or MESOPLUME when the plume passed the sampler arcs. The MSPUFF plume time of arrival, however, was better predicted than that of MESOPUFF and MESOPLUME in two of the three tables because the 850 mb winds tend to be larger than the 1500 m layer-averaged winds.

5.5.4. MESOPUFF II

The MESOPUFF II model contains a methodology for predicting plume dispersion and ground-level concentrations that may be considered as an extension or improvement of that in MESOPUFF. Plume transport occurs by means of a wind field generated by MESOPAC II. This meteorological preprocessor also provides gridded values of mixing height, stability and other micrometeorological variables. To be discussed first is the predictive accuracy of the model for plume trajectories and its implication for the accuracy of the plume model MESOPUFF II.

The accuracy of the MESOPAC II wind field prediction was determined by examining the accuracy of the predicted plume trajectories, as listed in Table 5-8 for SRP. In the prediction of wind fields for the Oklahoma and SRP data bases, MESOPAC II differs from MESOPAC in two key ways:

(a) MESOPAC II utilizes a two-layer wind field as compared to the single-layer wind field of MESOPAC, and

(b) MESOPAC II utilizes surface wind directions to help adjust the direction of the layer-averaged wind (lower layer).

The fact that the predicted plume often follows the Ekman spiral as it disperses vertically is supported by the statistics in Table 5-8, where the model predicted clockwise rotated plumes in 20 cases, but counterclockwise rotated plumes in only 10 cases. If a model shows both directions of rotation, that may reflect the naturally poor spatial and temporal resolution of the input wind data. If the clockwise direction occurs most of the time, it suggests an effect of the Ekman spiral, as discussed above. The 28 cases of correct plume direction compared favorably with the other models, with only MSPUFF predicting markedly fewer and MTDDIS predicting markedly more cases of correct plume orientation. For plume direction, the predictions of MESOPAC II appear to have the same accuracy as those predicted by MESOPAC. Both models showed a clear increase in clockwise rotated plumes as compared to counterclockwise rotated plumes.

For MESOPUFF II, categories (7) and (8) in Table 5-8 indicated that the rate of lateral plume spread is more accurately predicted than for MESOPUFF. Whereas the predicted plumes from MESOPUFF were never too wide (either about right or too narrow), those from MESOPUFF II were more balanced between too wide and too narrow (8 too wide versus 10 too narrow) for the SRP data. The main source of this difference is the use in MESOPUFF II of the user-input distance of only 10 km at which a transition takes place from dispersion rates based on the Turner curves to rates based on Heffter's curves. Since spreading rates implied by the Heffter curves are higher than those based on the Turner curves at large distances from the source, allowing the former to take effect closer to the source leads to considerably more spreading. The slight excess of too little spreading (5) over too much spreading (6), suggests that 10 km may still be too far from the source to make the transition, or that other scale considerations need to be included in specifying the lateral spreading rate.

Results shown in Table 5-7 do not contradict this point even though MESOPUFF II has the largest maximum concentrations of any of the models for the SRP runs. These very large concentrations occur at points near the source at times when the source is emitting krypton-85. However, beyond 20 km (the distance to the nearest sampler), the maximum concentration predictions of

MESOPUFF II are usually smaller than those of the other models.

The predictive results for the two Oklahoma releases yield a different comparison. The predictions of MESOPUFF (driven by MESOPAC) are uniformly more accurate than those of MESOPUFF II (driven by MESOPAC II). As Table 5-6 illustrates, maximum concentrations over the grid for MESOPUFF II were larger than those for MESOPUFF, both for the first few hours after release and during peak convective mixing hours; those for MESOPUFF were more nearly equal to the maximum observed concentrations.

The plume predicted by MESOPUFF II for the Oklahoma data sets seemed spatially more compact. After the first release, the plume passed to the east of the appropriate samplers, moved too slowly, arrived late and remained for 1-2 hours after the samplers ceased indicating tracer material above background. On the contrary, the MESOPUFF plume passed through the active part of both arcs, arrived on time at the 100 km arc, and had approached the 600 km arc more closely than MESOPUFF II during the averaging period when the samplers were first activated and measured nonzero plume concentrations. It appears that the predictive advantages that MESOPUFF II and MESOPAC II possessed over MESOPUFF and MESOPAC for SRP became disadvantages at Oklahoma. The meteorological conditions for the Oklahoma experiment were markedly different than those for the SRP experiment. The latter data cases represented a variety of conditions from four seasons, whereas the former represented only hot summertime conditions with steady wind flow from the south or southwest at most locations for most of the experimental periods. The maximum concentrations listed in Table 5-7 show lower values predicted by MESOPUFF II at distances beyond 100 km than those predicted by MESOPUFF. This behavior further confirms that the dispersion formulation of MESOPUFF II embodies more rapid plume spreading beyond 100 km than does the formulation in MESOPUFF.

However, in drawing preliminary conclusions on the relative merits of the theoretical formulations of the two models, it must be remembered that their meteorological preprocessors have sufficiently large theoretical differences and that considerable variability in predictive behavior can be expected. In MESOPAC, which provides the wind field and mixing heights for MESOPUFF, the wind field is simply interpolated from the layer-averaged winds at rawinsonde locations to grid locations, whereas in MESOPAC II a more detailed weighting scheme is used involving both the layer-averaged winds and surface winds from reporting stations. In MESOPAC, the mixing heights are obtained as the maximum of the convective and mechanical values computed without the use of measured cloud cover and ceiling height data. (Constant typical values are assumed for cloud cover and ceiling height.) In MESOPAC II, cloud cover and ceiling height data are used in connection with micrometeorological variables in obtaining both mixing heights and stability classes. Resulting values of stability and mixing height at grid points often differ substantially, affecting the predicted rates of plume spreading. In MESOPUFF II, unlike MESOPUFF, a two-layer wind field is also used, with puffs in the upper layer following a different trajectory than puffs in the mixed layer. When mixing heights rise in MESOPUFF II, either as a function of space or time of day, puffs in the upper layer can influence ground concentration values once again.

Our conclusion is that the results of this limited comparative study cannot discriminate between the effects of these differences in model formulation. Further and more detailed study would be needed to determine whether each of these differences produces a net improvement or net worsening of model predictive performance.

5.5.5. MTDDIS

The MTDDIS model uses considerably different theoretical assumptions than the four Gaussian puff models considered above. Because it is formulated only for elevated releases, it was applied solely to the Savannah River Plant cases. It contains its own meteorological preprocessor, which employs only wind data at the release height, rather than layer-averaged winds or two-layer winds. Its mixing heights are computed by a methodology that begins with NWS mixing height values from a single location in the experimental region. Other differences are noted in Section 2. The MTDDIS theoretical formulation is sufficiently distinct from the other models that it becomes more difficult to separate the effects of a single assumption in comparing the model's predictions with those of the other puff models. The important characteristics of the model/data comparisons of the MTDDIS model with the Savannah River Plant data base are (see Tables 5-7 and 5-8):

(a) plume trajectories appear to be the most accurately predicted of the seven models compared,

(b) plume concentrations tend to be overpredicted, and

(c) the model tends to make slight overpredictions of horizontal spreading.

A major distinction between the MTDDIS predictions and those of the other models is that MTDDIS overpredicts spreading whereas the other models underpredict spreading (except RTM-II for Oklahoma). The expected theoretical causes of these physical characteristics of the MTDDIS predictions are:

(a) Plume trajectories are well predicted by MTDDIS due to the use of wind directions that are representative of the effective height of release of the source at SRP. Other models employ wind directions taken from the 850 mb level, or computed as an integrated mixing layer average, or computed as an average to 1500 m. Apparently, for the modeling of the relatively short mesoscale distances at SRP (less than 150 km), wind directions at the effective height of release for industrial stack applications (not ground release applications) may be appropriate. Predictions for much longer distances (i.e., longer transport times) may require wind directions that represent puff movement at higher effective elevations from the ground than are used by MTDDIS.

(b) Plume concentrations tend to be overpredicted. All eight models exhibit this trend for both Oklahoma and SRP data sets, except RTM-II for Oklahoma. This is apparently a symptom that the vertical and horizontal mixing are not properly handled in any of the models. This overprediction occurs despite a considerable range of assumptions for modeling dispersion. For the dispersion methodology used in MTDDIS, as for each of the other approaches, there seems to be a need for better determination of the dispersion coefficients.

(c) The slight overprediction of horizontal spreading is a very interesting feature of the MTDDIS predictions. Two possible causes may be at work here. First, the use of the horizontal dispersion coefficient formulation taken from Heffter[31] applied directly from the release point undoubtedly adds more lateral spreading to the plume than the Turner curves (at least for the neutral and stable classes). Second, the use of hourly wind directions and speeds taken from the effective height of release provides, in general, a wider variation in transport time and direction of travel than winds with directions taken from the 850 mb level, or directions obtained from averaging over the mixing height or a fixed height, such as 1500 m. These greater variations lead to the transport of puffs to a wider range of directions, and cause the plume material to be spread out more along the direction of travel. In essence, then, the Heffter formulas appear to provide a greater spreading due to small-scale processes, and the use of wind directions at the effective height of release appears to provide greater plume meander and longitudinal stretching.

In summary, the MTDDIS predicted plumes cover a larger ground area than do those of the other models, and show a slight overprediction of lateral spreading when compared with the data.

A logical question arises then as to the accuracy of the mixing height algorithm since an overprediction of both horizontal spreading and 10-hour averaged concentrations occurs at ground level. It appears that mixing heights may be underpredicted based on simple conservation of mass principles. A detailed examination of this possibility was beyond the scope of the project. A re-examination of the choice of roughness height was made to determine whether values that were too small were used, producing a small mixing height. It was found that the MTDDIS computer runs had used values between 10 and 30 cm which were, indeed, characteristic of the land use in the modeled region.

5.5.6. ARRPA

Because the National Weather Service Boundary Layer Model wind field predictions were not available for the time period of the Savannah River Plant experiment, the ARRPA model could only be run for the Oklahoma experiment. Comments here on the model's predictive behavior are based on Tables 5-3 to 5-6. Evaluation of the results in these tables and of the statistical measures presented in Section 4 reveals the following characteristics of the ARRPA predictions:

(a) the BLM wind field provides plume transport that is quite accurate. The ARRPA plume arrives slightly later than the observed plume (generally by one averaging period). On the other hand, the ARRPA plume transport is more rapid than any of the predictions of the other six models tested at Oklahoma, and

(b) the ARRPA plume predictions are characterized by overprediction of concentrations and underprediction of lateral spreading.

These features characterize both ARRPA runs for the Oklahoma data base. Table 5-6 reveals that the ARRPA predictions for maximum concentrations were larger than the predictions of all of the models except MSPUFF and RADM, and tended to decrease less rapidly with downwind distance than those of MSPUFF. The same characteristic features become evident also in Tables 5-3 to 5-5. The predicted azimuthal angle of plume passage through both arcs compares favorably with the observed azimuthal angle of passage in all three tables. Plume widths are also close to the observed widths at the 100 km arc, but are too narrow at the 600 km arc. The transport times predicted by ARRPA are also close to the observed times.

The wind field is quite well predicted based on an examination of the location of the concentration distribution. Specific comparisons of wind components with upper air or surface data have not been carried out since such a comparison, although valuable, was beyond the scope of the present effort. The BLM model would be expected to do well in these Oklahoma cases since there was a well-defined pressure gradient present. The BLM model has difficulty in cases of light and variable winds[52], common in the summer season. In such situations, the surface topography becomes important, and the 80 km x 80 km grid provides too little spatial resolution.

The well-predicted plume transport for the Oklahoma runs is, however, surprising considering that a ground-level release was used in the experiment. The BLM model has difficulty resolving winds in the lower layers of the boundary layer[52]. The vertical velocities in the surface layer are essentially negligible. Apparently, the richer set of upper air and surface stations along with the more sophisticated physical treatment of the meteorological interpolation (based on conservation equations) explains the greater wind field prediction accuracy. It should be kept in mind that this evaluation has been only with two cases, and that further evaluation of the BLM model over a wide range of meteorological conditions (with ARRPA) is required before more conclusive comments can be made about the BLM wind field model.

The cause of the overprediction of concentrations and the underprediction of lateral spreading is likely due to the choice of K_y for the lateral dispersion coefficient. This coefficient, dependent on stability class, is used in the dispersion of plumes whose σ_y is greater than 1000 m. At the time step when σ_y becomes greater than 1000 m, a matching is made between Turner's curves and Gifford's Lagrangian theory to determine the K_y value in Gifford's theory. (In Gifford's theory, the lateral dispersion is a function of time and K_y.) In this way, a continuous plume width is maintained during transition between the two theories of lateral spreading. The value of K_y determined by this process was found to be on the lower end of the range of values found in the literature for dispersion in the mesoscale.[31] Moreover, a sensitivity analysis carried out by TVA indicated that model predictions were very sensitive to the choice of this parameter.[53] A choice of the higher values (based solely on measured values in the literature and not on a matching process) should lead to greater lateral spreading and reduced peak concentrations. This hypothesis as to the cause of model/data discrepancies is likely to be correct, but remains to be confirmed by means of actual computer simulations.

The following discussion is offered to explain the physical reason for the underprediction of lateral spreading. The Turner curves were developed from small-scale diffusion experiments, with the dispersion relationships generated from them often extrapolated to larger distances. On the other hand, the scale of the field of turbulence is quite different in the mesoscale with eddies involved in the mixing being kilometers in size, considering that plume widths are in that size range. It is conceivable then that the matching procedure does not provide a K_y that represents the true scales of turbulence that are occurring at these large distances.

It is expected that the conversion from Turner's curves to the Gifford Lagrangian theory has occurred early in the dispersion due to the presence of strongly convective conditions on the hot summer days of the releases. It is possible that the σ_y value became greater than 1000 m and the matching took place after the first hour of transport. Further numerical experiments with K_y including a rerun of the two Oklahoma cases, printing out the location and details of the matching (stability class, downwind distance, σ_y, time, etc.), would be very instructive.

If K_y is the problem, one possible improvement might be to continue the matching procedure as before, while adding a scale-dependency to K_y. Such a dependency would allow K_y to increase as the plume and the scales of turbulence which affect it grow.

It is interesting that recent validation tests of ARRPA carried out by TVA[54] suggest a tendency for underprediction of plume width and an overestimate of plume vertical extent. This validation work covered downwind distances only from about 10 to 70 km. An examination of plume cross-sectional areas in that study revealed no significant problem, undoubtedly because the underprediction of width and the overprediction of plume vertical extent offset each other. However, for larger distances downwind, the underprediction of lateral spreading continues (as shown here with the Oklahoma runs), whereas the overprediction of vertical spreading does not since it is limited by the mixing height. As a result, that compensation does not occur beyond these larger distances and the bias in the lateral spreading becomes more evident in the predictions. It is expected that this underprediction of lateral spreading is the main cause of the overprediction of concentrations at the ground.

Finally, the ARRPA model was one of four evaluated recently with the CAPTEX data sets.[54] In these seven data sets, the models were tested for their predictive capability out to distances of about 1000 km. The ARRPA model revealed a very significant underprediction of lateral spreading at the ground and a significant overprediction of plume concentrations at ground level. Consequently, the systematic behaviors observed in the current study are consistent with that longer distance testing.

5.5.7. RADM

One of the most obvious features characterizing the predictions of RADM for the Oklahoma data base and, except for MSPUFF, is its tendency to overpredict plume ground-level concentrations. This characteristic was observed with both the Oklahoma and SRP data bases. RADM predicts the highest maximum concentration values (over the whole grid) of any model within the first 100 km of the source (see Tables 5-6). In fact, even the MSPUFF peak predictions eventually fall below those of RADM during the first Oklahoma case. Its maximum concencentration predictions are the sixth highest of the seven models (on the average) for the Oklahoma data sets, exceeded only by MESOPLUME. Tables 5-3 to 5-5 also confirm this behavior in that plume widths on passage through the sampler arcs are among the smallest. When the RADM predicted plume passes the 600 km arc after the first release, the width is a factor of 4 smaller than the observed, and a factor of 2 below the next-smallest predicted width (by MSPUFF). Table 5-7 confirms this tendency to overestimate ground concentrations in that RADM predicts the second-highest maximum concentration values over the entire grid at SRP. (It was explained in Section 5.5.4. above that MESOPUFF II predicts very large near-source concentrations while the source is emitting, which causes it to predict the largest maximum concentrations much of the time at SRP.) Two reasons can be identified that are responsible for the predictive behavior of RADM described above.

(a) The eddy diffusivities used in the RADM plume dispersion calculations may be too low. Previous experience with the model at Dames & Moore, Inc, has only involved the modeling of plume dispersion in regions not exceeding 100 km square with a plume whose maximum lateral extent seldom exceeded 10 km. It is within this region where the RADM formulation of lateral and vertical spreading using the lateral and vertical diffusivities, K_H and K_V, agrees roughly with the Turner curves. However, it is known that at some distance less than 100 km from the source, one has to begin using growth rates of plume dimensions that take account of the effect of larger-scale eddies. One such formulation used by the other models is that of Heffter. If one continues to use the Turner curves to mesoscale distances, too little dispersion will result at large distances, as observed in the predictions of RADM. The experience with the MESOPUFF models in this study suggests that the transition region is more likely to be less than 10 km rather than 100 km or more. In addition, an increase in the diffusivities with scale may be needed in predicting plume dispersion for many hundreds of kilometers.

(b) For distances exceeding 100 km, it may be necessary to select the parcel mass so that more parcels are used to represent the plume. In these computer runs, the parcel mass was adjusted so that for each case at least 5000 parcels were present at some time during the run. The 5000 parcel figure was adopted in accord with the RADM user's manual guidelines; however, for mesoscale distances a larger number may be needed. Using too few parcels can lead to artificially high concentrations and an underestimation of the lateral plume extent. This behavior is due to the small number of widely dispersed parcels, which reduces the likelihood that a sampler will be impacted by one. (Of course, with more parcels, each one has a smaller mass.)

Concerning the positioning of the predicted plume, the RADM plume follows a trajectory determined by the wind field from MESOPAC II. Its trajectory was seen to agree with plume trajectories predicted by the RTM-II model (at Oklahoma) and the MESOPUFF II model. The winds predicted by MSPACK (for MSPUFF) are regularly less accurate than those of either the MESOPAC model (for MESOPUFF and MESOPLUME) or the MESOPAC-II model (for MESOPUFF II, RADM and RTM-II at Oklahoma). However, both MESOPAC models produce winds that are not as accurate in direction or speed as those from the BLM model (for ARRPA).

In the RADM model runs, parcels were released every 10 minutes, which is well within the guidelines offered in the user's manual. For some runs, marked differences were observed between the RADM model predictions and those from MESOPUFF II. This is interesting because both employed the meteorological preprocessor MESOPAC II. For case 13-A at SRP (July 15, 1977 from 0900 to 1900 GMT), RADM predicts a compact plume to the northwest of the source, extending about 70 km downwind, whereas MESOPUFF II predicts a much longer plume extending to the west-southwest. However, during the next averaging period for case 13-B, the ground level plume regions for the two models roughly coincide, extending to the southwest. Such differences emphasize the combined effect of differences in plume diffusion formulations and methods of computing concentration averages. The effects of these different factors are difficult to separate. Case 13-A illustrates another interesting point. During that averaging period, only two samplers, one 20 km from the source and one 50 km from the source, recorded plume concentrations significantly above background. Both models predict zero concentration at the farther sampler, and both predict a concentration within a factor of 2.0 of the observed concentration at the nearer one. However, the plume patterns are quite different. Even the statistical measures paired in space and time will be affected equally by these predictions, because the SRP sampler network does not provide enough spatial resolution to exhibit the difference in plume patterns. This case clearly shows the benefit of augmenting the AMS statistics with other methods for comparing predicted plume patterns.

The method of computing plume concentrations from parcel location and "size" may be responsible for some of these large prediction differences between RADM and MESOPUFF II. The horizontal and vertical σ's are used in RADM to define a rectangular "box" around each parcel, and the fractional overlap of each parcel's "box" with a user-defined receptor volume (also rectangular) is used to assign part of that parcel's mass as a contribution to the receptor concentration. (The parcel's volume is constrained in all three dimensions to be less than or equal to that of the receptor.) This interesting method differs markedly from the use of a Gaussian distribution of plume material to compute concentrations from puffs, and may partially explain the wide predictive differences between RADM and MESOPUFF II. With the present determination of its diffusion coefficients, RADM exhibits less predictive accuracy than does MESOPUFF II.

Finally, from Table 5-8 for the 65 Savannah River cases, one can draw several conclusions concerning the performance of RADM. The assignment of plume trajectories to direction categories yields 28 correct trajectories, which places RADM among the four most accurate models. When the direction of the trajectory is found to be offset from the observed trajectory (31 times), a clockwise rotation is twice as frequent as a counterclockwise rotation. As expected, these figures are almost identical to those for MESOPUFF II. However, because of the interplay of the model's Lagrangian random walk formulation, the stability-dependent diffusion coefficients and the concentration averaging method, the model appears to predict much narrower plumes with less lateral spread than does MESOPUFF II. The RADM model never predicts a plume wider than has been observed, and predicts a too narrow plume in 22 cases. On the other hand, MESOPUFF II predicts a plume that is too narrow only 10 times and a plume that is too wide 8 times, suggesting an approximately correct spreading rate. The reasons for the prediction of too-narrow plumes and too-high ground concencentrations by the RADM model for the SRP data base are the same as discussed in (a) and (b) above.

5.5.8. RTM-II

Of the six models that were applied to both data bases, RTM-II showed major predictive differences between the data bases. These differences arise from the differences in model application between the SRP and the Oklahoma data bases. The two differences in model application were the following:

(a) The meteorological data were provided from MESOPAC for the SRP predictions and from MESOPAC II for the Oklahoma predictions. Thus, trajectories are predicted for the plumes at SRP that match those for MESOPUFF and MESOPLUME, whereas at Oklahoma, the predicted trajectories match those for MESOPUFF II and RADM.

(b) The model developers selected values for the coefficient η, defined as the coefficient that multiplies the magnitude of the velocity field deformation tensor in the formulation of the gridded horizontal diffusion coefficient, K_H. Separate η values were selected by the model developers for the Oklahoma and SRP data base runs. They also selected different upper and lower bounds for this diffusion coefficient for each experiment. Since scale considerations enter into the choice of K_H, differences in these values between experiments are, a priori, reasonable. In RTM-II the η coefficient is a crucial user input that requires considerable engineering judgment.

5.5.8.1. RTM-II Features Common to Both the Oklahoma and the SRP Data Bases

The RTM-II model predictions are subject to large day-night variations in surface concentrations as the mixing layer follows its diurnal cycle. Since the mass of pollutant present is spread out over the mixing layer, mass conservation implies that significant decreases in the layer depth can cause large increases in surface concentrations, rather than the normal decrease with time due to diffusion. Also, significant increases in mixing height can cause decreases in ground concentrations that are more rapid than diffusion alone would cause. For the SRP cases studied, the averaging periods (10 hour) had been selected to represent primarily either daytime or nightime conditions, and the mixing heights did not show large diurnal variations within these periods. The long averaging times also tend to mask time-dependent effects in the observed concentrations. At Oklahoma, however, the first release was at 1900 GMT and sampling took place until 0300 GMT on the second day following, thus including both daytime and nighttime conditions. Table 5-9 shows the changes in the predicted MESOPAC II mixing height over this period for the location at the center of the predicted plume. During the period from 1100 to 1700 GMT, the predicted plume concentrations drop sharply, which seems to be explained by the rapid increase in mixing height and the spreading out of the available mass over a larger volume. However, the drop in concentration exceeds that which can be explained by the degree of increase in predicted mixing height. Upon examination, one other factor is probably responsible. In RTM-II, when the wind field shows divergence, mass that is being carried in the upper layer is transferred to the lower layer. The concept is that the divergent mass flow must come from air aloft. On the other hand, when the wind field is convergent, mass is transferred from the mixing layer to the upper layer. MESOPAC II does not compute a divergence-free wind field. When the wind field is convergent or divergent, the plume predicted by RTM-II loses mass to the upper layer or gains mass from it. In fact, during the period when the RTM-II concentrations are dropping so quickly, a front is approaching from the northeast, which normally produces

Table 5-9. Mixing height at plume center (rounded to 100 m) predicted by MESOPAC II for the first release at Oklahoma.

Day	Time	Mixing Height	Grid Location
July 08	1500 LST (2100 GMT)	2000 m	(15,11)
July 08	1900 LST (0100 GMT)	1100 m	(15,11)
July 08	2300 LST (0500 GMT)	600 m	(15,11)
July 09	0300 LST (0900 GMT)	600 m	(18,20)
July 09	0700 LST (1300 GMT)	400 m	(19,22)
July 09	1100 LST (1700 GMT)	1200 m	(26,31)
July 09	1500 LST (2100 GMT)	1300 m	(41,34)

convergent winds in the mixing layer. This behavior of the RTM-II model may be physically correct under the particular meteorology that was present after the time of plume passage (as determined from the measurements) through the 600 km arc. However, since there were no additional samplers active beyond 600 km, this conjecture cannot be verified by the data. A further investigation of how much of the physical convergence in the wind field caused by this approaching front is actually being predicted by MESOPAC II is beyond the scope of this evaluation.

5.5.8.2. RTM-II Features Specific to the Oklahoma Data Base

The plume trajectory predicted by RTM-II during the first release (July 8, 1980 at 1900 GMT to July 10, 1980 at 0200 GMT) stays eastward of the observed plume location and does not transport rapidly enough to arrive at the 600 km arc when the tracer was first observed on July 9, 1980 at 0200 GMT. However, there is a predicted wind speed gradient that is quite strong near this arc (from MESOPAC II); prior errors in wind speed prediction become amplified in this region. The predicted plume actually sweeps far to the east and exits the region moving east-southeast, whereas the data suggest a plume that should exit moving northward approximately in the center of the grid. The plume also arrives too late by more than an hour at the 100 km arc. During the second case, the predicted plume arrives at the 100 km arc at the correct time, but it passes 20-40 km to the west of samplers recording concentrations above ambient. It is also slow to leave this arc, producing sizable concentrations at samplers up to 2 hours after the data shows none. The apparent slow movement may be caused by too-rapid spreading, rather than errors in the wind speeds predicted by MESOPAC II. The directional problems do stem from methodologies in MESOPAC II.

For both cases, RTM-II predictions show extremely rapid spreading as illustrated in the isopleth graphs. Consistent with its rapid spreading, its peak concentrations, (listed in Table 5-6) are below those of any other model on an average over the 21 time periods. (The plumes predicted by the MSPUFF and ARRPA models yield zero maximum concentrations during the last two periods on July 9 because they exit the grid too soon.) This behavior of RTM-II described above is undoubtedly due to values for η (thus the horizontal diffusivity, K_H) and the minimum horizontal diffusivity, $(K_H)_{min}$, that are too large. (No value of K_H computed from the wind field deformation was found to be as large as the specified maximum value in these Oklahoma runs, but the minimum value was often assigned to grid points.) Refinements in the procedure for selecting η and $(K_H)_{min}$ as a function of scale seem desirable to improve the predictive accuracy of RTM-II. From Tables 5-3 to 5-5, one can see that the predicted widths on plume passage through the sampler arcs equal or exceed the observed widths, a feature found only in the predictions of RTM-II, and only for the Oklahoma experiment. The arrival of the predicted plume at the 600 km arc is late by about 8 hours, a feature common to the MESOPAC II-based-models.

5.5.8.3. RTM-II Features Specific to the Savannah River Plant Data Base

Plume trajectories in this case follow those for MESOPUFF and MESOPLUME. In Table 5-8, categories (1)-(4) characterize the plume trajectory and categories (5)-(6) describe rate of plume spreading relative to the data (with 65 minus the sum of categories (5)-(8) giving the number of predicted plumes that had approximately correct widths). For nearly half of the cases (31), the direction of the predicted plume is appropriate. There are slightly more plumes rotated clockwise than rotated counterclockwise. One would expect layer-averaged winds to show more clockwise rotation if, indeed, the surface wind direction were the correct one to use in predicting ground concentrations at these distances, as appears from the performance of the MTDDIS model. The RTM-II model also shows some complete misses, such as 6 plumes that were directed nearly opposite to the observed one, and the 4 cases of all zero predictions when some sampler recorded the presence of the plume. Plume spreading comes from only two factors in this model. First, if the wind were to change direction significantly during the averaging period, the plume would appear wider than diffusion alone would predict, because the plume would sweep across more grid points as its transport direction changes. Since the MESOPUFF model should also show this type of spreading, we can largely rule it out as a factor. The second factor is the magnitude of the horizontal diffusion coefficient. The tendency of plumes to be too narrow indicates an η coefficient that is somewhat too small. In this model, stability class values at the grid points are not used to influence the horizontal diffusion coefficient. It may be physically unrealistic to rely entirely on the deformation of the wind field to obtain the magnitude of the horizontal diffusion coefficient. The fact that the spreading of the plume for the Savannah River Plant data sets is too small and at Oklahoma it is much too rapid, underscores the need to develop a more accurate way of choosing these parameters for RTM-II once the grid size and scale of the calculation are defined. Despite the fact that the RTM-II model predicted plumes that spread too little with respect to the observed plumes, they spread more rapidly than those of any other models. As a result, the maximum predicted concentrations over the entire grid, as listed in Table 5-7, are less than those of any of the other models on an average over the 65 sampling periods.

SECTION 6

SUMMARY AND CONCLUSIONS

6.1. OVERVIEW

Eight short-term long-range transport models have been evaluated using two data bases representing different transport distances and averaging times. The Oklahoma data base contained two data sets for which the source was located at the ground and the emissions were 3-hour continuous releases (one for each experiment) of perfluorocarbon tracer. The concentration data encompassed detailed spatial measurements of the tracer at the 100 km and 600 km arcs with ground-level measurements made over a total of 21 time-averaging periods. The measured data were taken as 45-minute averages at the 100 km arc and 3-hour time averages at the 600 km arc. The very fine spatial resolution at 100 km and 600 km arcs provided a unique opportunity to evaluate the accuracy of the models in predicting the time of travel of the observed and predicted plume to these arcs along with the location, spreading, and pointwise concentration values of the predicted and observed plumes as they crossed those arcs. An interesting feature of this data base is the presence of a strong nocturnal jet that transported the plume very rapidly downwind.

The Savannah River Plant data base contained 15 data sets representing spatial measurements over distances of 28 to 144 km. There were 13 fixed samplers placed about a 62-m stack emitting a tracer gas, krypton-85. Measurements were available over several 10-hour averaging periods for each data set. These fifteen episodes complement the Oklahoma data sets in that they involve records of 2-5 days, but at a fewer number of fixed samplers. In spite of the differences in character of the data bases used in the evaluation, the performance of the models was generally consistent between the data bases.

Several methods were used to evaluate the models:

(a) statistical comparisons using the American Meteorological Society (AMS) statistics. These statistics employ pointwise pairs of predicted and observed concentrations,

(b) graphical comparisons of predicted/observed pairs including scatter plots, frequency histograms of residuals, cumulative frequency plots, etc. for each model for each data base,

(c) quantitative comparisons of the predicted and observed patterns for the Oklahoma data sets in order to evaluate their similarity. Compared were the peak concentrations, the angular offset of the pattern centerlines, and the time lag/lead of the predicted patterns from the observed, and

(d) qualitative comparisons of the ground-level concentration patterns predicted by the models and observed data for each averaging period. Concentration isopleths were presented for Oklahoma and sketches of the ground pattern were presented for SRP.

The results of these comparisons were largely consistent and depict the model performance described below.

6.2. GENERAL PERFORMANCE FEATURES OF THE MODELS

The key feature of the ground-level plume patterns predicted by the models is that they are frequently offset from the data by as much as 20-45 degrees. The treatment of the meteorological data within the meteorological preprocessor employed by a long-range transport model is the primary factor in the offset of the predicted and observed patterns. Often, the uncertainty in plume spreading is of secondary importance as compared to the uncertainties in characterizing the wind field. Due to the relatively short mesoscale distances involved in this model evaluation study (most of the data used in the evaluation were taken at distances less than 150 km), the accuracy of the wind field prediction (especially the initial direction of the plume) was crucial to the accuracy of the plume model predictions.

The BLM model (preprocessor to ARRPA) and the MTDDIS meteorological preprocessor appear to provide the most accurate trajectories of all eight models. However, the BLM model testing was carried out only for the two Oklahoma cases; more testing with that model is required before definitive conclusions can be reached. The MTDDIS preprocessor employs wind directions estimated at the effective height of release based on surface data. The use of low elevations for wind direction determination (rather than the 850 mb level, or the integrated average over 1500 m, or an average to the mixing height) appear to have provided good estimates for predicted trajectories at ground level in the distance range tested here, i.e. less than 150 km. However, the MTDDIS model was not tested with the Oklahoma data base since the model was considered by its developers to be inapplicable to that site due to the presence of a ground-level source. Definitive conclusions about the predictive accuracy of the MTDDIS-predicted trajectories also should await testing with data having greater spatial and temporal resolution and longer distances of transport than available from the SRP data base.

As noted above, the second most important cause of model/data discrepancies is the treatment of plume spreading in the models. This spreading may be viewed as superimposed about the trajectory given by the predicted wind field. Except for MTDDIS (at SRP) and RTM-II (at Oklahoma), the models underpredict horizontal spreading at both sites. In addition, the evaluation of predicted and observed ground-level concentrations revealed that each of the models (except RTM-II for SRP) overpredicts ground-level plume concentrations. (An approximate conservation of mass principle at ground level appears to be operational here; the vertical dimension is probably not as significant considering that after a short distance downwind, full mixing throughout the mixed layer has taken place.) The main cause of the underprediction of lateral spreading appears to be the often-used assumption that lateral spreading is given by the Turner curves for distances downwind out to 50-100 km. The models affected by that assumption are MESOPUFF, MESOPLUME, MSPUFF, ARRPA, and RADM. The more correct distance is about 10 km.

6.3. QUANTITATIVE MEASURES OF MODEL PERFORMANCE

The AMS statistics were applied to the predicted/observed pairs for the models at both sites. Statistical results are generally consistent between the two sites. The spatial and temporal offsets of the predicted and observed plumes lead to predicted concentrations that correlate poorly with concentrations observed at the same time and place. On average, all models overpredict (pairing in space and time) at both sites except for RTM-II which underpredicted only at Oklahoma. The largest overprediction is by RADM at both sites. Values for standard deviation of residuals, root mean square error, and absolute average residual are larger than average observed concentrations for all eight models. The largest values for all three measures occur with RADM. Predicted concentrations also are found to correlate poorly with concentrations observed at the same time and place. Pearson's correlation ranges among the models from -0.14 (MESOPUFF II) to 0.68 (MESOPUFF) for Oklahoma and -0.02 (MSPUFF) to 0.35 (MTDDIS) for SRP. Variance ratios are consistently less than 1.0 for all models except RTM-II.

On the other hand, statistical comparisons of the peak values predicted by the models are significantly better. Statistics for highest concentration by event (unpaired by location) reveal an overprediction by all models except RTM-II at Oklahoma. Statistics for highest concentration at each station (unpaired in time) also reveal overprediction by all models except MESOPUFF and RTM-II at Oklahoma. The best statistics are achieved through unpairing in both space and time of the highest 25 predictions and highest 25 observations. Based on the results presented in Tables 4-9 and 4-10, the ratios of the averages of the highest 25 predictions to the highest 25 observations are:

Model	Oklahoma	SRP
MESOPUFF	1.4	1.4
MESOPLUME	1.8	1.4
MSPUFF	3.5	1.3
MESOPUFF II	1.2	1.1
MTDDIS	N/A	1.4
ARRPA	1.8	N/A
RADM	3.3	4.7
RTM-II	0.7	1.0

Clearly, then, as more unpairing is accomplished, the statistical results improve significantly. Notice that in the above table, most models tend to overpredict peak concentrations. However, the advantage of an overprediction of the peak values in a regulatory setting must be weighed against (a) the spatial offsets that occur between the predicted and observed patterns, and (b) the underprediction of lateral spreading that has been shown to occur for these models (exceptions are MTDDIS for SRP and RTM-II for Oklahoma).

A pattern comparison method was used at Oklahoma to quantify the differences in ground-level configurations of the observed and predicted plumes. The large numbers of samplers at these arcs made quantification possible to a limited degree. At the 600 km arc for example, it was found that:

(a) the predicted plume transport time lags the observed time of transport for all models with the slowest plume (RADM) arriving at least 9 hours late,

(b) all models underpredict plume widths as the peak concentration passed over the arc. The range is 50 km width (RADM) to 180 km width (RTM-II) as compared to a measured width of 200 km,

(c) an angular offset of the location of the predicted and observed peaks on the arc is usually found, and that offset varied from 11° (MESOPUFF II) to 24° (MESOPUFF, MESOPLUME, RTM-II, and RADM). For this first Oklahoma experiment, the observed peak crossed at 10° clockwise from north.

Mixed results occurred for the 100 km arc. Some predicted plumes (from MESOPUFF, MESOPLUME, and MSPUFF) arrive earlier than the observed plume at that 100 km distance. The remaining models (MESOPUFF II, ARRPA, RADM, RTM-II) show actual plume arrival later than predicted plume passaage.

6.4. SPECIFIC PERFORMANCE FEATURES OF INDIVIDUAL MODELS

6.4.1. MESOPUFF

The MESOPUFF model frequently reveals errors of 20 to 40 degrees in the direction of the predicted plume (predictions usually clockwise relative to observed) along with an underprediction in plume spread. Model errors are most likely due to (a) the single-layer nature of the wind field analysis based on twice-daily rawinsonde data used in MESOPAC (inadequately treating the diurnal cycles of mixing depth and vertical shear in the horizontal wind, as well as not treating meandering), and (b) an underprediction of small-scale processes; i.e., too small σ_y. Concentrations are generally overpredicted as compared with the other models and with the data.

6.4.2. MESOPLUME

The MESOPLUME model employs the same wind field as does MESOPUFF; as a result, the same relative ground positioning and times of arrival are found with both models. However, the models show significantly different ground concentration predictions. In general, the MESOPLUME concentration predictions are up to a factor of two larger than the MESOPUFF predictions, with a smaller lateral spreading as compared to MESOPUFF. Compared with the data, MESOPLUME generally overpredicts concentrations and underpredicts widths at ground level.

6.4.3. MSPUFF

This model predicts plume trajectories that are usually oriented clockwise of the trajectories predicted by MESOPAC-based-models. As noted above, the trajectories predicted by MESOPAC are themselves rotated clockwise with respect to the data. The explanation for MSPUFF relates to the use of 850 mb winds in the meteorological preprocessor MSPACK (as compared with the averaging of winds over a 1500 m depth, as done in MESOPAC). The concentrations at ground level from MSPACK are usually higher than for MESOPUFF and higher than the observed concentrations as well. Predicted plume widths are generally more narrow than indicated from the data. Possible interpretations of the causes of model/data discrepancies relate to the treatment of the mixing height (treatment of turbulent dissipation of convection for the convective mixing height in the Benkley-Schulman method and/or the use of surface winds in modeling the mechanical mixing height).

6.4.4. MESOPUFF II

The MESOPUFF II model also tends to overpredict plume concentrations and has a tendency to underestimate plume widths. At the SRP site, the MESOPUFF II model predicted about an equal number of plumes that were too wide or too narrow, whereas at the Oklahoma site an underprediction of spreading occurred. The model tends to overpredict concentrations at both sites. The use of 10 km as the transition distance between the use of the Turner curves and Heffter formula for lateral dispersion definitely improved concentration predictions over MESOPUFF. The two-layer wind field model appeared to be an advantage at SRP but a disdavantage at Oklahoma. The advantage of having two wind fields that are essentially uncoupled may be in providing a better treatment of nocturnal wind shear. At Oklahoma, a single-layer wind field model appeared to perform better due to the fairly uniform flow there from the south or southwest.

6.4.5. MTDDIS

This model was only applicable to the SRP data sets. The model provided the best prediction of plume trajectories for this data base, yet revealed a slight overprediction in horizontal spreading and an overprediction in predicted concentrations. The use of a single-layer wind field with the wind direction taken at the effective height of release apparently provides good transport directions for the puffs at least for the short mesoscale distances in the SRP data base. The wider range in wind directions that occurs at such a low level (at the effective height of release from the 62-m stack as compared, say, to directions at the 850 mb level) enhance lateral spreading. The use of the Heffter formula for lateral dispersion from the source also provides more lateral spreading under most conditions than in the other models that use the Turner curves. These latter two effects are the likely causes of the overprediction of horizontal spreading. The cause of the overprediction in the concentration predictions may be the result of the mixing height model.

6.4.6. ARRPA

This model was applied only to the Oklahoma data sets. The BLM wind field model provides the most accurate trajectories of the competing wind field models at Oklahoma. The physical model used in interpolating the surface and upper air data appears to have been quite good for these two cases in which strong pressure gradients dominated the flow. In spite of the good trajectories, the lateral dispersion was underpredicted and the plume concentrations were overpredicted. The formulation for K_y in the Gifford Lagrangian model for lateral dispersion (σ_y in the range between 1000 and 6000 m) appears to have been underestimated.

6.4.7. RADM

This model uses the MESOPAC II meteorological preprocessor and thus has the same features in terms of trajectories as does MESOPUFF II. The horizontal spreading is underpredicted and the ground-level concentrations are overpredicted. The use of K_H and K_V as horizontal and vertical dispersion coefficients (tied closely to the Turner curves) are likely causes of this behavior. For mesoscale applications, a new choice of coefficients K_H and K_V that leads to plume growth rates consistent with the presence of the larger-scale eddies should be considered for future model improvement. It is also possible that the method of computing ground concentrations from parcel masses and locations is producing some of the occasional wide differences in the predicted plume pattern relative to MESOPUFF II for the same cases.

6.4.8. RTM-II

The predictions of this model were quite different between the SRP and Oklahoma sites due to (a) the different choice of η (the constant that multiplies the deformation of the wind field to produce the horizontal diffusivity) for the lateral dispersion coefficient for both sites, and (b) the different choice of meteorological preprocessors. At SRP, the model overpredicted plume concentration and underestimated plume spreading. At Oklahoma, the model overpredicted plume spreading and underpredicted plume concentrations. Considerable judgment is required in the choice of this lateral spreading coefficient. The grid spacing used in the plume modeling as well as the regional scale of the modeling problem are both inputs into this judgment process.

6.5. OPTIONS AND PARAMETERS

Finally, as alluded to earlier, most of these models (MESOPUFF, MESOPLUME, MSPUFF, MESOPUFF II, RADM, and RTM-II) require special judgment in their application. The choice of options is often crucial to the prediction accuracy. For RTM-II, the choice of the constant η determines how much horizontal spreading will occur in the predictions. For MESOPUFF II, the choice of the transition distance between the use of the Turner curves and the Heffter formulas also determines the amount of horizontal spreading that will occur in a particular application. In addition, the choice of wind field option in many of the preprocessors is equally if not more important. As a result, the accuracy of these models in future applications depends to a large degree on intelligent user choices for such key parameters and options.

REFERENCES

1. United States Environmental Protection Agency. "Guideline on air quality models", Report EPA-450/2-78-027. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, 1978.

2. Federal Register, "Guideline on air quality models," 45:20157-20158, March 27,1980.

3. D.G. Fox, "Judging air quality model performance. A summary of the AMS workshop on dispersion model performance. Woods Hole, Mass., 8-11 September 1980. "Bulletin American Meteorological Society, 62:599-609, May 1981.

4. G.J. Ferber, K. Telegadas, J. Heffter, C. Dickson, R. Dietz, and P. Krey "Demonstration of a long-range atmospheric tracer system using perfluorocarbons," Air Resources Laboratories, NOAA Technical Memorandum ERL ARL-101, Silver Spring, Maryland. April 1981.

5. K. Telegadas, G. Ferber, R. Draxler, M. Pendergast, A. Boni, J. Hughes, and J. Gray. "Measured weekly and twice-daily krypton-85 surface air concentrations within 150 km of the Savannah River Plant (March 1975 through September 1977) -final report, " NOAA Technical Memorandum ERL ARL-80. Air Resources Laboratories. Silver Spring, Maryland. January 1980.

6. C.W. Benkley and A. Bass. "User's guide to MESOPUFF (Mesoscale Puff) Model," Environmental Research and Technology, Inc., Concord, Massachusetts. September 1979.

7. C.W. Benkley and A. Bass. "Development of mesoscale air quality simulation models. Volume 6. User's guide to MESOPAC (Mesoscale Meteorology Package)," EPA-600/7-79-XXX, Environmental Research and Technology, Inc., Concord, Massachusetts. September 1979.

8. C.W. Benkley and A. Bass, "Development of mesoscale air quality simulation models. Volume 2. User's guide to MESOPLUME (mesoscale plume segment) model," EPA-600/7-79-XXX, Environmental Research and Technology, Inc., Concord, Massachusetts. September 1979.

9. M.R. Schock and S.F. Weber. "Modification of the MSPUFF model (a version of MESOPUFF) treatment of plume dispersion," Division of Environmental Waste Management and Research. North Dakota State Department of Health. Bismarck, North Dakota. December 1982.

10. M.R. Schock and S.F. Weber. "MSPUFF mesoscale air quality computer modeling system," Division of Environmental Waste Management and Research. North Dakota State Department of Health. Bismarck, North Dakota. July 1, 1984.

11. J.S. Scire, F.W. Lurmann, A. Bass, and S.R. Hanna. "Development of the MESOPUFF II dispersion model," Environmental Research and Technology, Inc. Concord, Massachusetts. Prepared for Office of Research and Development. U.S. Environmental Protection Agency. Research Triangle Park, North Carolina. 1984.

12. J.S. Scire, F.W. Lurmann, A. Bass, and S.R. Hanna. "User's guide to MESOPUFF II model and related processor programs," Environmental Research and Technology, Inc., Concord, Massachusetts. Prepared for Office of Research and Development. U.S. Environmental Protection Agency. Research Triangle Park, North Carolina. 1984.

13. J.S. Scire and F.W. Lurmann. "Development of the MESOPUFF II dispersion model." Paper presented at the AMS Sixth Symposium on Turbulence and Diffusion. Boston, Massachusetts. March 22-25, 1983.

14. I.T. Wang, T.L. Waldron, "User's guide for MTDDIS, mesoscale transport, diffusion, and dispersion model for industrial sources," Report EMSC6062.1UG(R2), Rockwell International, Inc., Newbury Park, California. December 1980.

15. S.F. Mueller and R.J. Valente. "Meteorological data preprocessing manual for the Air Resources Regional Pollution Assessment Model (generic version)," Office of Natural Resources, Air Quality Branch. Tennessee Valley Authority. Muscle Shoals, Alabama. September 1983.

16. S.F. Mueller, R.J. Valente, T.L. Crawford, A.L. Sparks, and L.L. Gautney, Jr. "Description of the Air Resources Regional Pollution Assessment (ARRPA) Model." Tennessee Valley Authority. Air Quality Branch. Muscle Shoals, Alabama. September 1983.

17. D.I. Austin, A.W. Bealer, W.R. Goodin. "Random-walk advection and dispersion model (RADM)," Dames & Moore, Inc., Los Angeles, California. December 1981.

18. M. Yocke, R. Morris, M. Liu. "Revised user's guide to the Regional Transport Model," Publication No. 82120. Systems Applications, Inc., San Rafael, California. April 1982.

19. A.J. Policastro, M. Wastag. L. Coke, R.A. Carhart, and W.E. Dunn, "Evaluation of eight short-term long-range transport models with field data, Task I Report: Preparation of input for eight long-range transport models", Prepared by Argonne National Laboratory, Argonne, Illinois and the University of Illinois at Chicago and Champaign-Urbana, for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, January 1985.

20. A.J. Policastro, M. Wastag, L. Coke, R.A. Carhart, and W.E. Dunn, "Evaluation of eight short-term long-range transport models with field data, ADDENDUM to Task I Report: Preparation of input for eight long-range transport models", Prepared by Argonne National Laboratory, Argonne, Illinois and the University of Illinois at Chicago and Champaign-Urbana, for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, May 1985.

21. A.J. Policastro, M. Wastag, L. Coke, R.A. Carhart, and W.E. Dunn, "Evaluation of eight short-term long-range transport models with field data, Task II Report: Preparation of test cases and proposed statistical/graphical evaluation methods (revised)", Prepared by Argonne National Laboratory, Argonne, Illinois and the University of Illinois at Chicago and Champaign-Urbana, for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, July 1985.

22. A.J. Policastro, M. Wastag. L. Coke, R.A. Carhart, and W.E. Dunn, "Evaluation of eight short-term long-range transport models with field data, ADDENDUM to Task II Report: Preparation of test cases and proposed statistical/graphical evaluation methods (revised)", Prepared by Argonne National Laboratory, Argonne, Illinois and the University of Illinois at Chicago and Champaign-Urbana, for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, February 1985.

23. A.J. Policastro, M. Wastag, J.D. Shannon, R.A. Carhart, and W.E. Dunn, "Evaluation of two short-term long range transport models with field data". Transactions, APCA Specialty Conference. The Meteorology of Acid Deposition. Perry Samson, Ed. October 6-9, 1983. Hartford, Connecticut.

24. C.W. Benkley and L.L. Schulman, "Estimating hourly mixing depths from historical meteorological data", Journal of Applied Meteorology, 18:772-780, 1979.

25. D.B. Turner, "Workbook of atmospheric dispersion estimates", U.S. Department of H.E.W., Public Health Service, Publ. 999-AP-26, 1970.

26. P.R. Maul, "Atmospheric transport of sulfur compound pollutants", Central Electricity Generating Bureau MID/SSD/80/0026/R, Nottingham, England, 1980.

27. A. Venkatram, "Estimating the Monin-Obukhov length in the stable boundary layer for dispersion calculations", Boundary Layer Meteorology, 19:481-485, 1980.

28. A. Venkatram, "Estimation of turbulence velocity scales in the stable and unstable boundary layer for dispersion applications", IN: Eleventh NATO-CCSM International Technical Meeting on Air Pollution Modeling and its Application, p. 54-56, 1980.

29. D.B. Turner, "A diffusion model for an urban area", J. Applied Meteorology, 3:83-91, 1964.

30. D. Golder, "Relations among stability parameters in the surface layer", Boundary Layer Meteorology, 3:46-58, 1972.

31. J.L. Heffter, "The variations of horizontal diffusion parameters with time for travel periods of one hour of longer", Journal of Applied Meteorology, 4:153-156, 1965.

32. R.R. Draxler, "Modeling the results of two recent mesoscale dispersion experiments", Atmospheric Environment, 13:1523-1533, 1979.

33. J. Smagorinsky, "General circulation experiments with the primitive equations: I. The basic experiment", Mon. Wea. Rev., Vol 91: 99-164, 1963.

34. B.C. Scott, "Parameterization of sulfate removal by precipitation", J. Applied Meteorology, 17:1375-1389, 1978.

35. J.M. Hales and S.L. Sutter, "Solubility of sulfur dioxide in water at low concentrations," Atmospheric Environment, 7:997-1101, 1973.

36. T.E. Pierce and D.B. Turner, "User's guide for MPTER", Environmental Sciences Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1980.

37. F.A. Gifford, "Horizontal diffusion in the atmosphere: a Lagrangian-dynamical theory", Atmospheric Environment, 16:505-512, 1982.

38. G.J. McRae, W.R. Goodin and J.H. Seinfeld, "Development of a second-generation mathematical model for urban pollution", Atmospheric Environment, 16:679-696, 1982.

39. C.C. Shir, "A preliminary numerical study of atmospheric turbulent flows in the idealized planetary boundary layer", J. Atmos. Sci., 30:1327-1339, 1973.

40. J. O'Brien, "On the vertical structure of the eddy exchange coefficient in the planetary boundary layer", J. Atmos. Sci., 27:1213-1215, 1970.

41. J.C. Wyngaard, "Modeling of the planetary boundary layer - extension to the stable case", Bound. Layer Met., 9:441-460, 1975.

42. Environmental Protection Agency, "User's manual for single source (CRSTER) model", EPA-450/2-77-013, Office of Air Quality Planning and Standards,U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1977.

43. J.L. Heffter, "Air Resources Laboratories Atmospheric Transport and Dispersion Model (ARL-ATAD)", NOAA Technical Memorandum ERL ARL-81, Air Resources Laboratories, Silver Spring, Maryland, February 1980.

44. R.R. Draxler of Air Resources Laboratory, NOAA. Personal communication with Dr. I.T. Wang of Rockwell International Corp., Newbury Park, California, 1979.

45. P.R. Slawson, J.H. Coleman and J.W. Frey, "Natural draft cooling tower plume behavior at Paradise steam plant: Part II", Division of Environmental Planning, Tennessee Valley Authority, Publication No. TVA/EP-78/01, February 1978.

46. M.E. Smith, "Recommended guide for the prediction of the dispersion of airbourne effluents", American Society of Mechanical Engineers, 1968.

47. J.P. Boris and D.L. Book, "Flux corrected transport — I. SHASTA, a fluid transport algorithm that works", J. Comput. Phys., Vol II: 38-69, 1973.

48. R.J. Londergan, D.H. Minott, D.J. Wackter, T. Kincaid, and L. Boutata, "Evaluation of rural air quality simulation models." TRC Environmental Consultants, Inc. Report EPA-450/4-83-003. October 1982.

49. R.J. Londergan, D.H. Minott, D.J. Wackter, and R.R. Fizz, "Evaluation of urban air quality simulation models." TRC Environmental Consultants, Inc, EPA-450/4-83-020, July 1983.

50. D. Wackter and R. Londergan, "Evaluation of complex terrain air quality simulation models", TRC Environmental Consultants, EPA-450/4-84-017, June 1984.

51. C.S. Hirtzel and J.E. Quon, "Estimating precision of autocorrelated air quality measurements", IN: Summary of Proceedings of Environmetrics 81:200-201, 1981.

52. Mueller, S. of Tennessee Valley Authority. Personal Communication with A.J. Policastro, Argonne National Laboratory, March 1986.

53. Mueller, S.F. and L.M. Reisinger, "Evaluation of the air resources regional pollution (ARRPA) model", Office of Natural Resources Air Quality Branch, Tennessee Valley Authority, Muscle Shoals, Alabama, February 1986.

54. Shannon, J. of Argonne National Laboratory. Personal Communication with A.J. Policastro, Argonne National Laboratory, November 1985.

55. C.M. Shieh, M.L. Wesely, and B.B. Hicks, "A guide for estimating dry deposition velocities of sulfur over the eastern United States and surrounding regions", Argonne National Laboratory, ANL/RER-79-2, Argonne, Illinois, April 1972.